

FEASIBILITY OF USING CAPI/CATI INTERVIEWERS TO CONDUCT COGNITIVE INTERVIEWS DURING SURVEY FIELD TESTS

Paul A. Mullin, Research Triangle Institute
P. O. Box 12194, Research Triangle Park, NC 27709-2194

Key Words: Cognitive interview, Interviewers, Field Test

INTRODUCTION

Rapid survey development schedules and tight budgets sometimes preclude survey organizations from conducting cognitive laboratory pretests of data collection instruments prior to field testing. When cognitive pretests are conducted, the surveys are often sent to the field without an additional round of testing to confirm that question revisions improved the instrument and did not produce new problems. In this context, survey designers may be interested in the feasibility of conducting some form of cognitive pretesting as part of survey field test activities.

Laboratory Versus Field Test Settings

Controlled laboratory settings are commonly used for cognitive interviewing in order to reduce spurious response errors resulting from distraction and to provide an environment where respondents can concentrate on thinking aloud. In addition, the laboratory setting is usually convenient for trained methods staff. Nevertheless, researchers conduct cognitive interviews in a variety of settings such as at the respondents' homes or offices, public libraries, fast food restaurants, health clinics, etc.

Some consideration should be given to the possible advantages of conducting cognitive interviews in a setting that closely resembles the setting in which the instrument will be fielded. Along these lines, recent research suggests the advantages of conducting cognitive interviews of telephone survey instruments on the telephone (Beatty & Schechter, 1994; Schechter, Blair, & Vande Hey, 1996).

It is conceivable that conducting cognitive interviews under field test conditions could reveal context-specific question problems. For example, respondents may effectively handle questions with a high working memory load when the questions are presented in the quiet, focused environment of a laboratory setting. Under field conditions, however, numerous distractions can reduce the ability of the respondent to concentrate. As a result, respondent performance with such questions may suffer differentially in comparison to performance with other questions. This may suggest the need to incorporate simpler terms in the question or reduce the complexity of the question sentence structure. Such a

need may be missed in the laboratory setting.

Cognitive Interviewing Goals and Methods

Cognitive pretesting can take many different forms and utilize a variety of methods (Forsyth & Lessler, 1991). Cognitive interviewing usually involves asking the respondent to "think aloud" and respond to probe questions about the meaning of the questions or the strategies and processes by which the respondent arrived at his/her answers. Investigators use specific "directed" probe questions to confirm suspected problems that are identified in a professional review of the instrument. For example, a cognitive interviewer may ask "Did you interpret [term or phrase] to mean [definition A] or [definition B]?" Investigators use broader, "general" probe questions to discover problems which may be present. For example, a cognitive interviewer may ask "How did you arrive at your answer?" Once a question answering problem has been identified, an investigator may ask respondents to provide input on potential solutions to identified problems.

In sum, cognitive interviewing seeks to confirm suspected problems, discover unknown problems, and engineer solutions to problems that are identified. In general, an effective cognitive interviewer must adapt the probing to the issues that arise during the course of an interview. Similarly over the course of a series of cognitive interviews, the investigator may adapt the protocol to include additional directed probes aimed at confirming problems and determining solutions to problems that are discovered in earlier interviews.

Can Regular Survey Interviewers Conduct Cognitive Interviews?

Anecdotal reports from some survey organizations suggest that their past attempts to use regular interviewers for cognitive interviewing have been less than successful. Indeed, many survey methodologists will agree that regular interviewers are usually not ideal candidates for the role of cognitive interviewer. In pointing to the difficulty of training regular survey interviewers to conduct effective cognitive interviews, Willis (1994) contrasts the usual job role of interviewers with the requirements of cognitive interviewing. In their usual role, interviewers are asked to make questions work, operate at a fast pace, focus on question structure, and ask standardized questions. In contrast, a cognitive interviewer is asked to find flaws in the questions, work

deliberately, focus on question content questions, and formulate questions on the fly (or decide which fixed probes are appropriate in a given situation).

Nevertheless, the use of regular interviewers for cognitive interviewing may be pervasive in certain types of survey organizations. For example, Blair and Presser (1993) reported that two-thirds of academic survey research organizations use at least some of their regular survey interviewers to conduct cognitive interviews. Moreover, several published research studies demonstrate success in using regular survey interviewers to conduct cognitive interviews. An early report was provided by Belson (1981). In this study, interviewers were specially selected during the hiring process and received additional training in asking probe questions to identify question answering problems.

Oksenberg, Cannel, & Kalton, (1991) also reported success in using regular field test interviewers to collect cognitive data. These researchers embedded a variety of different types of probe questions in a questionnaire. For example, they used open-ended probes on comprehension and information retrieval. To limit to impact of the probe questions on the interview length, they create several versions of the instrument, each containing a subset of the probes.

The Bureau of the Census used closed-ended, post-interview, respondent debriefing questions to examine question comprehension by respondents in the Current Population Survey (Campanelli, Rothgeb, & Martin, 1991; DeMaio & Rothgeb, 1996). This included asking respondents to choose the appropriate response for a hypothetical person in response to the information presented in a vignette. The methods developed for this research were clearly crafted very carefully to fit interviewers' usual job requirements and to avoid interrupting the course of the survey interview. Yet, they yielded very useful cognitive data.

Consistent with the above findings, another line of research shows that interviewers can converse with respondents about question meaning to improve data quality (Schober & Conrad, 1997a, 1997b). While not directly involving cognitive interviewing, these results suggest that interviewers may be able to discern when a respondent is having difficulty with a question and can flexibly and appropriately discuss the question meaning.

The Difficulty of Pretesting in Field Tests

Given the published research findings, it is somewhat surprising that cognitive testing methodologies have not been employed more often in field test settings. One consideration is that the investment required for selectively hiring interviewers and/or training them to conduct cognitive interviews may be prohibitive given job turnover rates among

interviewers. Another factor may be the difficulty of incorporating pretesting activities in survey field tests. Tight field test schedules, the burden of starting a field operation, and heavy interviewer training workloads may make the inclusion cognitive pretesting activities problematic.

More likely, however, cognitive pretesting activities are not included in field tests because most organizations view this to be too late in the survey development process. By the time of the field test, a considerable investment has been made in programming the instrument, developing manuals, training interviewers, and preparing other aspects of the survey operations. It is difficult for most organizations to contemplate making major revisions to an instrument with these investments at stake.

Nevertheless, many question problems identified in cognitive pretesting may be improved or eliminated without substantial revisions to the instrument or through interviewer training. It seems reasonable that survey organizations would in at least some cases be able to schedule sufficient additional time for revisions after the field test. Moreover, many studies involve repeating data collections and revisions to the instruments may be made after the initial data collection cycle based on cognitive data collected in the field test. Cognitive data may be useful when one goal of the field test is to compare similar sets of questions for the purpose of determining which set will be used in the main study and which will be cut from the survey. Even when cognitive data cannot be used to refine the instruments, research analysts should find the validity assessment provided by pretesting useful when examining and interpreting the study data. This may help analysts to avoid drawing false research conclusions.

The Present Research Study

The goal of the present investigation was to investigate the utility of an easily implemented and inexpensive option for conducting cognitive testing in the context of a field test. An important focus was to develop procedures that did not require any special selection of interviewers and involved very limited interviewer training. Such an option might allow organizations to examine the performance of survey instruments using cognitive methods when limited funds are available or survey development schedules are too restrictive for standard laboratory testing prior to a field test. In addition, these procedures would not overburden interviewers attempting to learn their other field test duties, and would not require a substantial organizational investment in selecting and training interviewers.

METHODS

The present research study was conducted as part of the 1998 Field Test for *Child and Family Well Being Study* conducted by the Research Triangle Institute (RTI). The field test was conducted in three cities in the United States. Because the welfare programs and program names differ in the states where these cities are located, there was interest in testing the welfare component of the survey in all three cities. This posed some difficulty given time constraints, budgetary concerns, and the effort to arrange pretests in locations where RTI has no offices. The field test was viewed as an appropriate venue for addressing concerns about the instrumentation for the welfare component of the survey.

Content of the Welfare Component

The welfare section included questions asking respondents to provide a history of their welfare participation. This reporting was facilitated through the use of a life history calendar, which interviewers and respondents set up at the beginning of the interview. The section also included questions on the experiences of the respondent in the welfare program and their knowledge of the new welfare program rules.

An initial round of eight cognitive interviews was completed at RTI in Research Triangle Park, North Carolina. The interviews were conducted in a laboratory setting, and the questionnaire was revised in response to the findings. Based on these interviews, 16 questions were selected for cognitive testing in the field test. Some questions were selected to test revisions that had been made or because the limited laboratory testing left certain questions unanswered. Others questions were selected because of their importance in the welfare section.

Probing Methods

According to the study design, intensive training in cognitive interviewing was not provided to interviewers. Thus, it was unreasonable to assume that the interviewers would be very effective in follow-up probing. Instead the method relied largely on the interviewers' use of the scripted probes, although in several cases interviewers were given instructions for probing only. Both "directed" probes aimed at confirming suspected problems and general probes aimed at permitting the discovery of hitherto unknown problems were provided to interviewers. All probes were open-ended questions, and an attempt was made to examine both comprehension and recall issues.

A brief lead-in was scripted for the beginning of the welfare section of the survey. The content of this lead-in was generally typical of the introductory remarks made prior to cognitive interviews. The respondent was informed by the interviewer that they would be asked

additional questions to see what we could learn about their understanding of the questions and how they answered them. Respondents were told the goal of these additional questions was to improve the questionnaire.

In the CAI instrument, a "cognitive probe screen" was built into the instrument immediately following the screens for each of the 16 selected questions. Programming specifications for the cognitive probe screens were included with the main body of the interview, and the probe screens were constructed as part of the regular programming work.

Each probe screen had a label at the top which read "DISCUSS QUESTION" to alert the interviewers of to the type of question they would be asking. Scripted probe question and/or instructions for probing were provided below the label. Toward the bottom of the screen was a text entry area that was several lines long. Interviewers were instructed to enter responses to the probe questions in this area.

Interviewer Training

The welfare survey study field test was a particularly complex endeavor which involved interviews with several household members, psychological test administrations, mother-child interactional observations, and observational visits to day care centers. Interviewer training for the field test lasted a full 11 days. This intensive interviewer training schedule precluded any significant coverage of the cognitive interviewing procedures. This fit well with the study goals. Training on cognitive interviewing methods and the probe question screens lasted 20-30 minutes. In addition, the interviewer manual contained a four page section on the cognitive interviewing, including example probe screens.

RESULTS

Interviewer Reactions

Prior to the field test, the survey research staff held an orientation meeting for the field supervisors. This meeting involved a presentation of some of the data collection methods which would be used in the study, and included a brief discussion of the cognitive interviewing procedures. The field supervisors expressed considerable surprise and excitement that interviewers would actively assist in refining the instrument for the main study. A similar reaction was observed among the interviewers themselves during training. Some were pleased at the opportunity to avoid the frustrations of dealing with instrument problems which might be avoided with their input. One advantage to involving interviewers more actively in survey revisions may be the sense of empowerment it engenders among the interviewers. This may have significant benefits for their morale and performance.

Limitations in the Data Collected

Because of excessive length of the welfare study interviews and the overwhelming burden of the other survey activities, it was necessary for the survey director to cancel the cognitive testing procedures. Interviewers were instructed during and after training not to ask cognitive probe questions. The probe questions remained embedded in the instrument, however, and 28 of 42 interviewers recorded responses to the one or more of the probes in 100 of the 200 field test interviews.

As might be expected, many of the responses did not address the intended focus of the probe questions. Other entries were directly relevant, but did not provide sufficient information to answer the research questions. For example, many responses to a general probe about the meaning of a question consisted of a straightforward paraphrase of the question which did not reveal the respondents underlying understanding of the content. Naturally, these are the kinds of responses that a well-trained cognitive interviewer will follow-up on with additional tailored probe questions.

Despite the fact that many responses were not helpful, valuable information was collected about a number of the selected study questions. This included information on the comprehension of specific terms or entire questions, recall strategies, the use of prior instructions during answering, and follow-up information to explain apparently contradictory responses. Some illustrative findings are reviewed below.

Illustrations of Findings

Term Comprehension. Several directed probes were scripted to solicit the respondents' interpretation of specific terms. In one series of questions, respondents were instructed that we wanted to know about their participation in the welfare program that provides cash assistance for families, and were asked to give the name they used for this program (e.g., TANF, AFDC). They were then asked if they are currently receiving benefits from this program. Interviewers probed "What do you think we mean by "benefits"?" As suspected, many respondents very easily slipped into thinking about non-cash benefits, and a few apparently indicated they were receiving welfare benefits when they were receiving only Medicaid or food stamps.

Question Comprehension. Comprehension problems were also discovered using general probe questions. Respondents were asked a knowledge question about the dependence of welfare time limits on whether parents are working or not. The intention of the questions was to see if respondents knew that holding a job extends the length of time a person can stay on welfare. Interviewers probed "What does this question mean to you?" The

responses showed that some of the respondents interpreted the question to mean that you get cut off of welfare when you get a job. Depending on the income from the job, these respondent may be correct. However, the results indicated that they understood the question in a way that was not intended and that the question required revision.

Use of a Previous Instruction. Another of the cognitive probe questions was designed to ascertain whether respondents had followed a prior instruction when answering a survey question. Respondents were instructed at the start welfare history questions not to include welfare that their parents had received for the respondent as their child. Several questions later, respondents who were not receiving welfare currently were asked if they had ever received welfare. Interviewers probed whether or not respondents remembered the instruction. Of the 18 respondents who provided of relevant responses, 2 reported that they had included welfare that their parents had received.

Helpful Landmark Events. As part of collecting a welfare history, respondents were asked to report the start and end dates for their participation. Interviewers were instructed to probe on how the respondent recalled the dates they reported and whether recalling certain other events had helped them. Landmark events for welfare start dates included the birth of child, loss of spouse, and moving. Landmark events for welfare end dates included starting a job, family births/deaths, and finishing school. These data suggested the kinds of events that might be solicited when setting up the life history calendar in the beginning of the interview.

Sources of Contradiction. During a review of the survey, the research staff noted a sequence of questions in the welfare section for which a respondent could provide seemingly contradictory responses. Specifically, respondents might report that they were not required to work by the welfare program, but were not excused from the welfare rules. Alternatively, respondents might report that they were required to work, but had not worked even though they had not been excused from the rules. In the event that someone provided such a report, a cognitive probe screen instructing interviewers to determine how these reports might be justified was included. Responses to the probe indicated that people who were already working, people who were starting work in the near future, and people who were temporarily not following the rules because of an illness might provide these reports.

DISCUSSION

The present research demonstrated the utility of a easily implemented and inexpensive methodology for collecting cognitive interview data during survey field

tests. By this method, all regular field test interviewers are enlisted in the cognitive testing process. Open-ended cognitive probes and probing instructions are scripted by the researcher and built into the CAI instrument following selected test questions. The interviewers administer the probes as part of the regular interview and enter responses into their computers.

Valuable information on question problems and answering strategies was collected in the context of the field test with minimal effort and expenses. This included both the confirmation of suspected problems and the discovery of unanticipated problems. Information was collected on both comprehension issues and recall strategies.

Very little or no time was invested in training the interviewers about cognitive interviewing or about the specific issues of interest in the cognitive testing. Instead, the interviewers were provided with a brief orientation to the task during training and in their manual. Of course, the success of the present method depends heavily on the effectiveness of the cognitive protocol that is scripted for the interviewers by the researcher.

The large number of cognitive interviews that are conducted by the method suggests the possibility of quantifying the results to determine, for example, response error probabilities. However, this was not possible given the relatively high percentage of responses that were unrevealing. Unless interviewers are more fully trained in cognitive interviewing and have full knowledge of the issues being addressed in the cognitive testing, the present method will produce only qualitative findings. In actuality, the findings were fairly similar in nature to that which is normally obtained in laboratory testing involving small numbers of respondents.

Because of the lack of follow-up probing by interviewers, the method clearly relies on the collection of relatively large numbers of responses to produce meaningful results. The researcher must sort through the responses to find evidence of problems, and must make several inferences as to a) the source of the problem, b) whether the occurrences are representative of a problem that will occur with some frequency, and c) whether the problem can be fixed without difficulty. These inferences are very similar to the inferences that are typically made in cognitive laboratory pretests. Of course, the data available for inference-making in the laboratory setting may be more substantial depending on the quality of the cognitive interviewing staff.

Alternative approaches

One drawback of cognitive interviewing during field testing is the added length to the interview. Unlike the typical laboratory testing situation where only a part of a

regular interview may be administered, the full survey is administered in a field test. Moreover, field test interviews tend to be lengthier than is intended for the main study because of interviewers are inexperienced with the instrument, the instrument may be unrefined, and cuts to the instrument may be pending the field test results.

Some time savings might be obtained by audio recording the open-ended probe responses instead of having the interviewers enter them into the computer. However, the additional expense to make the recordings (e.g. audio tapes or disk space for digital recordings) and the additional labor to review or transcribe the audio records may be prohibitive.

The alternative approach provided by Oksenberg et al., (1991) involves testing only a limited set of questions in a given interview. If, however, the cognitive probes are built into the CAI instrument for the pretesting, additional programming to control when a given probe question is asked (i.e., based either on the interview or the interviewer) will be necessary.

Much of the expense of cognitive testing in the present study was related to programming the cognitive probe screens and removing them from the CAI instrument after field test. As an alternative, it may be possible to give interviewers a paper listing of the probe questions and add some type of marker to the question screens for which a probe is to be asked. A side benefit of this approach is that the original question screen will be available to the interviewer during probing. Interviewers might still enter the responses into their computer, for example, using the interviewer notes section that is regularly programmed into CAI instruments.

One improvement to the method that should be made relates to asking multiple probe questions after a given survey question. Some difficulty was encountered in the present study in sorting out the responses to different probes. If multiple probe questions are necessary, it should be possible to structure the task so that probe responses can be linked more directly to probe questions. This may involve asking the interviewers to identify the probe associated with the response in their notes (e.g., by assigning codes to the probes), or by dividing the notes spaces accordingly.

Final Thoughts

In conclusion, there are a number of options that may be used to adapt the method to the needs of a particular study and organization. The method undoubtedly cannot provide as complete or thorough an evaluation as laboratory research conducted by professional cognitive interviewers, but nevertheless can provide useful information for refining survey

instruments. When cognitive interviews are not possible due to schedule or budgetary constraints, or when additional information about the performance of an instrument is desired after laboratory pretesting activities are complete, the method employed in the present study can be helpful.

Interviewer focus groups are commonly used to collect information about operational difficulties in field tests, and frequently questionnaire issues are addressed as well. Yet, surprisingly, organizations do not provide any training for interviewers on how to evaluate the questions in an instrument. The benefits of providing a general orientation to interviewers about detecting question problems and actively involving them in the survey development process seems readily apparent. Interviewers jump at the opportunity to help refine a questionnaire. In reality, enhancing the feedback obtained from interviewers and increasing their job satisfaction may be accomplished with something as simple and basic as teaching them to ask general comprehension probes when they believe a question may be misunderstood.

ACKNOWLEDGMENTS

This research was funded by a professional services contract from the National Center for Health Statistics and by a subcontract to RTI from The Johns Hopkins University for the NSF grant-funded *Child and Family Well-being Study*. I am grateful to Andrew Cherlin, Allen Duffer, and Susan Schechter for their support and assistance.

REFERENCES

- Beatty, P., & Schechter, S. (1994). "An examination of mode effects in cognitive laboratory research" *Proceedings of the American Statistical Association, Section on Survey Methods Research* (pp. 1275-1280). Alexandria, VA: American Statistical Association.
- Belson, W. A. (1981). *The Design and Understanding of Survey Questions*. Aldershot, England: Gower.
- Blair, J. & Presser, S. (1993). "Survey procedures for conducting cognitive interviews to pretest questionnaires: A review of theory and practice." *Proceedings of the American Statistical Association, Section on Survey Methods Research* (pp. 1110-1115). Alexandria, VA: American Statistical Association.
- Campanelli, P., Rothgeb, J., & Martin, E. (1991). "The use of interviewer and respondent debriefing studies as a way to study response error in survey data." *The Statistician*, 40, 253-264.
- DeMaio, T. J. & Rothgeb, J. M. (1996). "Cognitive interviewing techniques: In the lab and in the field." In N. Schwarz and S. Sudman (Eds.), *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research* (pp. 177-195). San Francisco: Jossey-Bass Publishers.
- Forsyth, B. H. & Lessler, J. T. (1991). "Cognitive laboratory methods: A taxonomy." In P. B. Biemer et al. (Eds.), *Measurement Errors in Surveys*. New York: Wiley.
- Oksenberg, L., Cannell, C. & Kalton, G. (1991). "New strategies for pretesting survey questions." *Journal of Official Statistics*, 7, 349-365.
- Schechter, S., Beatty, P., & Vande Hey, J. (1996). "Conducting cognitive interviews to test self-administered and telephone Surveys: Which methods should we use?" *Proceedings of the American Statistical Association, Section on Survey Methods Research* (pp. 10-17). Alexandria, VA: American Statistical Association.
- Schober, M. F., & Conrad, F. G. (1997a). "Does conversational interviewing improve survey data quality beyond the laboratory?" *Proceedings of the American Statistical Association, Section on Survey Methods Research* (pp. 910-915). Alexandria, VA: American Statistical Association.
- Schober, M. F., & Conrad, F. G. (1997b). "Does conversational interviewing reduce survey measurement error?" *Public Opinion Quarterly*, 61, 576-602.
- Willis G. (1994). *Cognitive Interviewing and Questionnaire Design: A Training Manual*. National Center for Health Statistics: Cognitive Methods Staff, Working Paper No. 7.