

RESULTS FROM A COGNITIVE EXPLORATION OF THE 1993 AMERICAN HOUSING SURVEY

Dawn R. Von Thurn and Jeffrey C. Moore, U.S. Bureau of the Census¹

Dawn R. Von Thurn, Center for Survey Methods Research, Census Bureau, Washington, DC 20233-9150

KEY WORDS: Cognitive Interviews, Questionnaire Pretesting, Questionnaire Design

1. INTRODUCTION AND BACKGROUND

Since 1973, the American Housing Survey (AHS) has been conducted by the Census Bureau for the Department of Housing and Urban Development. The survey's primary purpose is to collect information about the size and composition of the U.S. housing inventory and to measure change over time.

Among many other key estimates, AHS data yield a housing quality index derived from questions about heating equipment, plumbing, kitchens, hallways and general upkeep -- questions intended to assess a housing unit's physical quality. For quite some time, the reliability of the index has posed problems to analysts. Large differences have been found between reinterview and original interview responses, and the general experience of the reinterview program has been inconsistent, unstable answers (Lessard and Bushery, 1993). In recent years, researchers have also detected unexplainable, statistically significant differences between computer assisted telephone interviewing and personal interviews with the index questions, differences which have continued to defy easy explanation (Meier and Bushery, 1994).

These problems, and the absence of obvious solutions to them, led to a call for exploratory research to identify potential causes and to suggest improvements to the question series. This paper describes the qualitative research program that was carried out. The centerpiece of this program was the use of cognitive interview methods to generate hypotheses about, and to suggest solutions to, respondent reporting problems in the AHS.

2. THE RESEARCH PROGRAM

The research program included two primary components, the first of which was a thorough review of the entire questionnaire. We convened an "expert panel" of current AHS staff to elicit their informed opinions regarding problems in the instrument. We also observed both computer-assisted telephone and

personal visit field interviewers as they administered the AHS instrument. We reviewed AHS item nonresponse and edit/imputation data. Finally, we (and others) critically reviewed the instrument from a general questionnaire design perspective.

Both complementing these initial investigations, and following up on the hypotheses they generated, was the primary research activity -- a series of 27 cognitive laboratory interviews to explore respondents' comprehension and interpretation of the survey questions. We conducted the interviews in two phases.

The initial phase of interviewing was driven by the assumption that for some key items -- the type of equipment used to heat the unit, in particular -- a respondent's knowledge would likely vary with tenure and structure type. Thus, we chose to focus our recruiting efforts in approximately equal measure on multi-unit renters and single-unit owners for the first phase of cognitive laboratory interviewing.

While this strategy proved effective for this one purpose, it had the unintended side effect of limiting the cognitive interview respondents almost exclusively to middle-class people living in middle-class homes; we failed to recruit any respondent whose residence could be considered to have significant physical problems. To rectify this, all respondents in the second phase of cognitive interviewing were residents of a local low-income, public housing complex.

For both phases of cognitive interviewing, the interview started with several practice questions to familiarize respondents with "concurrent think- aloud" procedures. In addition to this technique, we also employed question paraphrasings, qualitative timings, and various follow-up probes.

3. RESULTS

3.1 Strengths of the Cognitive Interview Method

A number of recent papers (e.g., Campanelli, et al., 1991; DeMaio, et al., 1993; Esposito, et al., 1991; Forsyth and Lessler, 1991; Jenkins and Dillman, 1993; Presser and Blair, in press; Tanur and Fienberg, 1992) have addressed the strengths and weaknesses of different pretesting methods. This literature has noted that a particular strength of the cognitive interview method is its ability to reveal the meanings respondents attribute to survey questions (or at least the meanings respondents care to verbalize to an interviewer). For example, Presser and Blair evaluate four pretesting methods, including cognitive interviews. They find that the bulk of the problems identified by the cognitive interview method are semantic in nature, typically involving a discrepancy between respondents' understanding of the meaning of a question and what the question designer intended. They also find cognitive interviews useful for identifying task problems -- such as difficulty recalling, formulating, or reporting an answer.

Our research offers numerous examples of the sort of semantic difficulties Presser and Blair refer to, particularly where ambiguous or ill-defined terms led respondents to widely varying interpretations of a survey question. For example, in this AHS question,

"In the last 3 months, was there any time when ALL the toilets in the home were not working?"

the words "not working" were clearly troublesome to our cognitive interview respondents. Some respondents used a narrow definition (the one the survey designers intended) and interpreted "not working" as "unusable" -- i.e., the toilet is stopped up and won't flush. Others defined "not working" much more broadly, as in "not working properly" -- e.g., the water is constantly running but the toilet is still usable. Because the words "not working" are ambiguous, respondents were forced to come up with their own definitions. Many of the respondents' definitions were quite legitimate, but simply did not match the survey designers' intent.

Other AHS questions yielded very similar and immediate lessons in the cognitive laboratory. For example, in the question

"Does the (house/apartment) have open cracks or holes in the inside walls or ceilings?"

the use of the word "open" as the single clue to the types of cracks or holes of interest to the survey also caused difficulties. One respondent, asked to explain a "yes" response, mentioned having holes in her walls from trying to hang curtains; another noted a cracked bathtub wall tile; a third reported hairline cracks in the wall by her front door. All of these examples of "open cracks or holes" are outside the intended scope of the question. Once again, ambiguous and poorly-defined terms forced respondents to create their own definitions, which, not surprisingly, varied from one person to another.

Respondent semantic problems can arise not just from ambiguous or ill-defined terms, but also when key words or phrases are buried or otherwise not attended to. For example,

"Has water leaked into your home from outdoors in the last 12 months? For example, has water leaked through the roof, basement, walls, closed windows, doors or skylights? Exclude plumbing or other inside leaks."

Read at leisure, it is clear that this question is intended to focus on water leaks coming into the home from the outside. However, in the cognitive interview setting it was just as clear that many respondents completely missed the inside/outside distinction. One respondent mentioned having a leaky kitchen sink, while another described her

overflowing toilet. When asked to rephrase this question, this latter respondent stated that it was asking whether she had had any leaks in her home in the last year.

After interpreting the meaning of a survey question, a respondent's next task is to come up with a response. Cognitive interviews can also aid in identifying the strategies respondents use when developing a response to a question, and this information can highlight another kind of respondent task problem. For example, a very problematic AHS question for many of our respondents was the following:

"Please look at this card.

What type of heating equipment is used MOST to heat the (house/apartment)?

- *A central warm-air furnace with air vents or ducts to the individual rooms?*
- *Steam or hot-water system with radiators OR other system using steam or hot water?*
- *Electric heat pump?*
- *Other built-in electric units permanently installed in wall, ceiling, or baseboard?*
- *Floor, wall, or other built-in, hot-air heater without ducts?*
- *Kerosene, gas, or oil room heater(s)?*
- *Portable electric heater(s)?*
- *Stove(s)?*
- *Fireplace(s) WITH inserts, that is, installed equipment designed to circulate more heat into the room?*
- *Fireplace(s) with NO inserts?*
- *Some other type of heating equipment?*
- *None?"*

The twelve response options presented a daunting task to our cognitive interview respondents, many of whom -- even the homeowners -- either did not know the information at all or knew very little about their heating equipment. Most, however, did struggle through the task and provided an answer to the question. By asking them to verbalize how they arrived at their response, we observed several examples of respondents basing what appeared on the surface to be a very solid answer on very little real information. Many respondents desperately searched the flashcard for any vaguely familiar term. For example, the term "baseboard" appears to have directed one respondent to choose the fourth option ("other built-in electric units"). This respondent's father had told her that she needed a special type of screw to hang curtains on her "baseboard" walls; thus, she concluded that she must also have baseboard heat.

Other respondents tried to answer the heating equipment question by using what they knew (or thought they knew) about the fuel used by the heating equipment. Several respondents, when reviewing the options, were clearly searching for the words "gas" or "electric" in order to respond to the question. One respondent expressly chose "electric heat pump" because it was the first option she encountered with the word "electric" in it. Closely observing and having respondents verbalize their response strategies made it very evident that highly suspect response strategies can yield quite legitimate-appearing answers.

3.2 Weaknesses of the Cognitive Interview Method

Our cognitive interviews identified numerous respondent semantic problems and respondent task problems in the AHS instrument; however, the respondent's production of a response is just one component of the survey process, and cognitive interviews are not well suited to identify errors arising from other important components of the process. Measurement problems attributable to the interviewer, for example, are typically not detectable by cognitive interviews.

The "heating equipment" question, discussed above, provides a good illustration. Because our research plan included observations and other qualitative methods, we have some information about how interviewers actually implement this difficult question -- in a word, idiosyncratically. No field or telephone interviewer we observed read the question as it is worded. In addition, interviewers use a number of different strategies to assist respondents. In personal visit interviews, some interviewers actually ask to see the heating equipment, and then the interviewer and the respondent, together, select the best response. This strategy is obviously not possible over the telephone, so telephone interviewers often use extensive verbal probes. Had our research only included cognitive interviews, we would not have uncovered this important measurement issue.

The validity of cognitive interview data presents a much more fundamental problem. Cognitive interviews are most useful to the extent that they elicit from a respondent reports about thought processes which accurately represent the "natural" processes that would have occurred during a standard survey interview, as the respondent heard and interpreted the question and arrived at a response. It may be impossible to make these inherently covert processes overt without introducing biases, but a skilled interviewer, using appropriate techniques, can perhaps minimize them. We opted to emphasize concurrent think-aloud probes over retrospective probes (although we used both techniques), under the assumption that getting respondents to think out loud as they were answering a question would yield a more accurate rendering of natural cognitive processes than would an after-the-fact recall of what they had been thinking at the time they were asked the question. There is no hard evidence to support this assumption, and it may be the case that trying in this way to make a normally internal process external, so it can be measured, almost necessarily changes the process. For example, respondents who engage in very little cognitive processing when answering survey questions may have felt compelled to think more thoroughly before responding to our questions, resulting in data that do not reflect these respondents' natural thought processes.

This sort of bias, while suspected to be an ever-present risk, is very difficult to detect. A more mundane cognitive interviewing problem, especially with concurrent think-aloud probes, is their all-too-frequent failure to elicit from respondents much information at all. Asking respondents to verbalize their thoughts is for many people neither a typical request nor an easy one to comply with -- in response to a question, people are socialized to simply provide an answer, not to explain what they are thinking. Some of our respondents had a very difficult time with the concurrent think-aloud probe to "tell me what you are thinking." Other researchers have reported similar experiences. There are several possible reasons for this phenomenon: respondents may not understand exactly why or how we want them to "think aloud," or may be suspicious of our motives; others may simply be uninterested in the survey topic or the interviewing experience in general; and it is also possible that interviewers lack the particular skills for drawing out such reports effectively.

Regardless of the reason, this phenomenon clearly affects the usefulness of cognitive interviews. In several of our interviews it was difficult to distinguish when a respondent did not have compelling thoughts from when he or she simply could not or did not articulate them; this left us uncertain whether the respondent actually understood the original question.

Hindsight suggests that we should have spent more time at the beginning of each session practicing the think-aloud procedure to make sure respondents fully understood what we were asking them to do. Our speculation, however, is that for some respondents virtually no practical amount of training would be sufficient. For these respondents, the researcher may want to emphasize other types of cognitive techniques. In our interviews, we often found that a paraphrasing probe -- such as "what is this question asking you?" -- would elicit substantial valuable information when concurrent think-aloud probes were not effective. Other techniques less dependent on verbal reports, such as card sorting, may also be useful in such situations.

4. DISCUSSION

In our experience investigating the AHS, we found cognitive interviews to be a valuable questionnaire design tool. When the goal is to understand why problems occur during the response process, cognitive interviews can yield important insights, which can then direct the questionnaire designer in revising survey questions. (For a further description of this research, and the specific proposed revisions to the AHS questionnaire, see Von Thurn, Jenkins, and Moore, 1994.)

However, cognitive interview methods are not without their weaknesses. Important types of survey measurement problems are not amenable to cognitive interview detection; the essential validity of the data produced is open to question; the skills of the cognitive interviewer are probably critical to its success; and some respondents may be unable to report their cognitive processes effectively. It is important to acknowledge these weaknesses so that steps can be taken to overcome them, perhaps the most important of which is to combine cognitive interviews with other methods of identifying questionnaire problems. Finally, of course, the real utility of cognitive interview methods in improving survey questionnaires can only be demonstrated by quantitative proof that the revisions suggested by the cognitive investigation actually yield more complete and more accurate survey reports.

NOTES

1. We thank our colleagues Cleo Jenkins and Bonnie Carver for their contributions to this research project, as well as John Cannon of the Census Bureau's Demographic Surveys Division for his support of the project and his helpful guidance throughout. The views expressed in this paper are the authors' and do not necessarily represent those of the U.S. Census Bureau.

REFERENCES

Campanelli, Pamela C., Elizabeth A. Martin, and Jennifer M. Rothgeb, "The Use of Respondent and Interviewer Debriefing Studies as a Way to Study Response Error in Survey Data," The Statistician, Volume 40, 1991, pp. 253-264.

DeMaio, Theresa, Nancy Mathiowetz, Jennifer Rothgeb, Mary Ellen Beach and Sharon Durant, Protocol for Pretesting Demographic Surveys at the Census Bureau, Report of the Pretesting Committee, Washington, DC: U.S. Bureau of the Census, June 28, 1993.

Esposito, James L., Pamela C. Campanelli, Jennifer Rothgeb, and Anne E. Polivka, "Determining Which Questions are Best: Methodologies for Evaluating Survey Questions," Proceedings of the American Statistical Association, Section on Survey Research Methods, 1991, pp. 46-55.

Forsyth, Barbara H. and Judith T. Lessler, "Cognitive Laboratory Methods: A Taxonomy," Chapter 20 (pp. 393-418) in Measurement Errors in Surveys, Paul P. Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz and Seymour Sudman, (eds.), John Wiley & Sons: New York, 1991.

Jenkins, Cleo R., and Don A. Dillman, "Combining Cognitive and Motivational Research Perspectives for the Design of Respondent-Friendly Self-Administered Questionnaires," paper presented at the Annual Meetings of the American Association for Public Opinion Research, May 1993 (revised).

Lessard, James M. and John M. Bushery, "Reinterview Report: Response Variance in the 1989 American Housing Survey--Metropolitan Sample," unpublished U.S. Bureau of the Census report, November 1993.

Meier, Andrea J. and John M. Bushery, "Response Errors in the American Housing Survey: Some Reasons Why They Happen," unpublished U.S. Bureau of the Census report, January 1994.

Presser, Stanley and Johnny Blair, "Survey Pretesting: Do Different Methods Produce Different Results?", Sociological Methodology (in press).

Tanur, Judith M., and Stephen E. Fienberg, "Cognitive Aspects of Surveys: Yesterday, Today, and Tomorrow," Journal of Official Statistics, Volume 8, Number 1, 1992, pp. 5-17.

Von Thurn, Dawn, Cleo Jenkins, and Jeff Moore, "Report on Qualitative Research on the Moderate Physical Problems Questions of the American Housing Survey," unpublished U.S. Bureau of the Census report, March 10, 1994.