Variation in Cognitive Interviewer Behavior-- Extent and Consequences

Paul Beatty, Susan Schechter, and Karen Whitaker, National Center for Health Statistics
Paul Beatty, NCHS, 6525 Belcrest Rd., Room 915, Hyattsville, MD 20782

Key Words: Cognitive interviewing, pretesting, probing, interviewer effects

Survey interviewers operate under well-defined constraints: they are generally instructed to read questions exactly as written, use only non-directive probes when necessary, and administer feedback to reinforce appropriate respondent behavior. Cognitive interviewers do not have a comparable set of rules that have been established and applied across organizations that rely on cognitive pretesting methods (Conrad and Blair, 1996; for general descriptions of cognitive interviewing methodology, see DeMaio and Rothgeb 1996; Willis 1994; Forsyth and Lessler 1991). We would expect cognitive interviewer behavior to be more varied than survey interviewer behavior-- a greater number of interviewer behaviors are considered acceptable, including directive probing and re-wording of questions. Cognitive interviews may be marked by frequent interruptions, requests for elaboration about the meanings of responses, and other deviations from the questionnaire.

Furthermore, specific application of these techniques may be left to individual interviewer discretion. This is further complicated by the fact that cognitive interviewing approaches evolve somewhat independently across organizations, reflecting the preferences of researchers in each (Beatty, in press). A set of definite standards of what should and should not occur in cognitive interviews has not been developed, evaluated, and systematically implemented. Thus, the potential for significant interviewer variation should be taken seriously.

Nevertheless, researchers have usually assumed that variation in cognitive interviewer behavior does not threaten the basic generalizability of findings to actual surveys (Beatty, Willis, and Schechter, 1996). It is important to explore the veracity of this assumption, especially given the recent finding that probing style has substantial influence on the precision of subject responses and reports of problems answering the questions (Beatty, Schechter and Whitaker, 1996). This paper explores in greater detail what cognitive interviewers actually do during interviews. A taxonomy of cognitive interviewer behaviors was developed from transcripts of cognitive interviews conducted at the National Center for Health Statistics (NCHS); this taxonomy distinguishes between various types of probes, feedback, and functional remarks. Interviews were coded using this taxonomy, and analysis reveals the extent of differences in the use of various behaviors across interviewers and across questions. We then consider the implications of our findings on the validity and generalizability of cognitive interview findings.

Method

In 1995, four members of the Cognitive Methods Staff of NCHS conducted 17 interviews using a "quality of life" module consisting of about nine questions. The interviewing protocol suggested probes that could be administered following each question, but did not rigidly constrain interviewer behavior or require that interviewers use every probe. Tape recordings of these interviews were transcribed.

The cognitive interviewers generally discussed one question at a time, using concurrent probing-- that is, the interviewer asked the survey question, the subject gave a response, and the interviewer immediately probed about that response before moving to the next question. Thus, these interviews essentially consisted of a series of short exchanges about individual questions. An example of one such exchange appears below:

(1) *Interviewer*: Would you say that in general your health is excellent, very good, good, fair, or poor?

(2) *Subject*: Excellent.

(3) *Interviewer*: Why do you say excellent?

(4) *Subject*: Because I eat all natural foods, and I'm a vegetarian. No meat or dairy products.

(5) *Interviewer*: No meat or dairy products. So was it easy for you to pick excellent?

(6) *Subject*: Oh yes. That was an easy question.

This exchange is first broken into lines, based on uninterrupted speech from one of the participants. Lines can consist of either individual or multiple utterances, meaning distinct questions or statements. For example, line 5 consists of two interviewer utterances: "No meat or dairy products," and "So was it easy for you to pick excellent." Line 3, on the other hand, consists of one utterance.

Cognitive interviewers generally read the initial question (line 1) as it was written in the questionnaire. Subsequent interviewer behavior, on lines 3 and 5 in this example, was more varied. While the interviewer was given a protocol with suggested probes, (s)he maintained considerable freedom to modify or ignore these suggestions, add other probes, and guide the discussion based on the content of each particular interview. Our analysis therefore focused on interviewer utterances that followed the subject's initial response to the question.

Within each exchange, subjects usually (but not always) answered the survey question. Some subjects answered immediately after hearing the question, but others digressed and offered commentary about the subject matter. Interviewers could address this situation in different ways: some might attempt to pin subjects down to an actual response as soon as possible, while others might allow subjects to digress, perhaps at the expense of obtaining a codeable answer to the question. Because of the potential variety of approaches, we also separated interviewer utterances that came before and after the subject's answer to the survey question. (If subjects never gave a codeable response, then all interviewer utterances were counted as appearing before the response.)

First Analysis— Amount of Interviewer Activity

One issue explored was the amount of utterances in cognitive interviews (temporarily ignoring the specifics of what interviewers did). In our initial analysis, we made simple counts of interviewer utterances and compared them across interviews (see Table 1, next column).

The first line of the table shows the mean utterances made by each interviewer, per interview-- in other words, the mean number of distinct statements made by each interviewer during probing for the entire questionnaire module (note again that this does not count the initial reading of the question). We can see that there is only slight variation across interviewers.

Table 1: Overall interviewer utterances

Interviewer[1]	1	2	3
Mean utterances			
Utterances per interview	37.3	34.6	38.7
Utterances per exchange	4.2	3.8	4.3
Before answer	2.1	1.6	1.8
After answer	2.1	2.2	2.5

The second line of Table 1 shows the mean utterances made by each interviewer per exchange— that is, the mean number of distinct interviewer statements that followed each survey question. This too did not vary much— interviewers generally made four utterances (usually probes) following each question.

In the third and fourth lines, we separated interviewer utterances that came before and after the answer to the survey question. There is some variety here— Interviewer 1 did more probing before the survey response than the other interviewers, and Interviewer 3 did more afterwards than the others. The proportions of probing before and after the survey response also differ somewhat.

This analysis is limited but raises an important question: are these differences due to interviewer behavior, or are they attributable to subjects' varying difficulty with the questions? The former case, in which some interviewers may be more focused on obtaining survey responses than others, has some important implications. One task of cognitive interviewers is to draw conclusions about subjects' difficulty in answering questions (Beatty, Schechter, and Whitaker, 1996); such judgments might be suspect if interviewers' attempts to obtain an answer vary a great deal. These data are insufficient to properly address this issue, although in the next section, we will explore in more detail what the cognitive interviewers actually did in this study.

Second Analysis— Type of Interviewer Probing

In studying our interview transcripts, we identified five distinct categories of interviewer utterances:

---

[1] A fourth interviewer who only conducted one interview was dropped from this analysis.

1. Traditional cognitive probes: Probes used to understand subjects' interpretation of terms, thought processes, and information that was taken into account when answering the question. These should appear commonly, since cognitive interviews are presumably designed to study these issues. Some widely held examples of cognitive probes include:

- "What were you thinking about while answering that?"
- "How did you come up with your answer?"
- "What does [term] mean to you?"
- "Was it easy to answer that-- did you know the answer or did you have to estimate?"

2. Confirmatory probes: Probes that ask subjects whether the information provided so far is correct. One form of confirmatory probe is mirroring— repeating back all or part of what a subject said verbatim as an implicit request for confirmation (see the second example below).

- "So, for the last 30 days you were only unhealthy for one day?"
- Subject: "My health is pretty terrible right now." Interviewer: "Pretty terrible?"

3. Expansive probes: Probes designed to obtain additional narrative information. Generally these probes lead to the heightened gathering of details and a more conversational exchange. For example:

- "Tell me more about your arthritis-- does that make it hard for you to get around?"
- "When did that [event] happen?"
- "Were you sick the entire day, or only part of it?"

4. Functional remarks: Probes that redirect the subject back to the original survey question, by repeating it or clarifying some aspect of it. These are essentially the opposite of expansive probes-- they are attempts to refocus the conversation on the specifics of the question.

- Repeat of survey question: "And how many days did you feel that way?"
- Clarifications: "Yes, I'm talking about how you felt in the last 30 days."

5. Feedback: Interviewer behaviors that are neither probes nor functional remarks, including traditional "survey feedback" about the subjects' performance on the survey, and also reactions to the actual substance of subjects' answers.

- Traditional: "Thanks, that's just the sort of information I'm looking for."
- Conversational: "I know what you mean, I feel the same way."

The interviewing protocol in this study (again, not a rigid script), suggested mostly "cognitive probes," as well as some "expansive probes."

Utterances from the transcripts were coded and grouped by interviewer; the results appear in Table 2:

Table 2: Mean Utterances by Interviewer

Interviewer[2]	1	2	3
Mean utterances per interview			
Cognitive	2.3	3.0	5.5
Confirmatory	17.3	8.8	9.2
Expansive	10.1	14.0	13.0
Functional	3.6	3.8	3.5
Feedback	4.0	5.0	7.5
Total per interview	37.3	34.6	38.7
Total utterances	n=261	n=173	n=155
Total interviews	7	5	4

There are noteworthy similarities and differences across interviewers. As mentioned previously, the total number of utterances per cognitive interview does not vary by much. Similarly, usage of functional remarks also appears to be stable across interviewers— presumably this activity is largely driven by the subject's answers rather than interviewer style.

The amount of actual cognitive probing also does not vary much across interviewers, but such utterances are surprisingly infrequent compared to other categories. Since cognitive interviews are ostensibly conducted to provide insights into cognitive processes (c.f. DeMaio and Rothgeb, 1996) we would expect more probes of this type. Actually, the majority of utterances for each interviewer may be characterized as either confirmatory probes or expansive probes. Still, preferences do vary across interviewers: Interviewer 1 was much more likely to use confirmatory probes, whereas Interviewers 2 and 3 relied more heavily on expansive probes. The remaining category, feedback, was the third-most

---

[2] The fourth interviewer who conducted only one interview was dropped from this analysis also.

1066

common type of utterance for all interviewers, falling behind confirmatory and expansive probes.

Interviewer utterances were also grouped by question, rather than interviewer, in order to assess whether the questions themselves were associated with variety of interviewers' behavior. One of the questions (Q1) was a standard subjective health assessment ("Would you say that in general your health is excellent, very good, good, fair, or poor?") Other questions on the test instrument asked subjects to report a number of days out of the past 30 days that a subjective health assessment applied to them. Table 3 looks at four of these questions in particular:

During the past 30 days, for how many days...

Q2) ... was your physical health not good?

Q10) ... did pain make it hard for you to do your usual activities?

Q13) ... have you felt you did not get enough rest or sleep

Q14) ... have you felt healthy and full of energy?

Table 3, below, shows the breakdown of utterances for each of these questions.

Table 3:   Mean Utterances by Question

Question	Q1	Q2	Q10	Q13	Q14
%Utterances by question					
Cognitive	17.1	12.5	14.3	0.0	12.0
Confirm	38.2	34.6	32.1	32.5	42.2
Expansive	27.6	25.0	35.7	50.0	31.3
Functional	7.9	12.5	14.3	6.3	6.0
Feedback	9.2	15.4	3.6	11.3	8.4
	100%	100%	100%	100%	100%
Total per question	n=76	n=104	n=28	n=80	n=83

The bottom row of the table, total utterances per question, gives an indication of how much interviewer activity is associated with each question. Note that relatively little interviewer activity followed administration of Q10. Most subjects quickly answered

"none"; consequently there was very little to probe or comment about. In contrast, Q2 generated many interviewer utterances. Conceptually, there is much to discuss regarding this question. Whereas most people do not experience incapacitating pain on a regular basis (as in Q10), many people have days of "bad physical health"-- which could involve illness, injury, chronic conditions and so on. Complex judgements may be involved in responding to these questions (Schechter, Beatty, and Willis, 1997), giving interviewers much to probe about. While the amount of interviewer utterances does not vary much across interviewers (Table 2), it does vary considerably by question. Total interviewer utterances may be a reasonable measure of the complexity involved in answering a particular question.

Note again that there are relatively few cognitive probes compared with confirmatory and expansive probes. No cognitive probes were administered following Q13 at all— interviewers seemed to prefer expansive probing (e.g., "How much sleep do you usually get?" which is perhaps more useful understanding subjects' answers than asking them directly about their thought processes). Also, functional remarks and feedback vary much more across questions than they do across interviewers.

Discussion

The primary goal of this study was to explore differences in behavior across cognitive interviewers when testing survey questions. We should recognize first of all that our observations are drawn from only one small cognitive interviewing study. Each study is unique— questionnaires that are tested vary widely, as do the objectives of cognitive interviewing rounds. Keeping this limitation in mind, we still believe that some of the results are informative.

Because interviewer behavior is not rigidly specified by their protocols, we expected to observe noticeable variety in what interviewers actually do, and this was the case. It is also somewhat reassuring to see that overall patterns of behavior were reasonably consistent— interviewers tended to use expansive and confirmatory probes most often, and all used similar amounts of feedback and functional remarks. Interviewers appear to follow a general pattern of activity, although there are some deviations— perhaps due to interviewer preferences, or specific circumstances arising in individual interviews, or both.

What is somewhat surprising is that traditional cognitive probes were used infrequently compared to other types of probes. Advocates of cognitive interviewing often claim

that the method is used to investigate cognitive processes while answering. If this is so, then interviewers obtain this data through a fairly indirect approach, and perhaps what has previously been considered as a "cognitive probe" should be re-assessed. A large proportion of interviewer behavior essentially confirms what subjects already said, or asks them to provide additional narrative information that elaborates upon their answers. We observed few attempts to directly access verbal reports of cognitive processes.

Our findings support Gerber and Wellens' (1997) assertion that cognitive interviewing as currently practiced has more to do with understanding the meanings of responses and questions than cognitive processes per se. This may be a profitable development: some researchers have long questioned whether subjects have reliable access to information about their cognitive processes (Nisbett and Wilson, 1977), and other critics of cognitive interviewing suggest that "cognitive probing" might actually alter the way subjects answer questions, creating unrealistic response conditions (c.f. Willis, 1994). The actual "confirm and expand" probing we observed is less obtrusive, and perhaps represents an important step in the evolution of cognitive interviewing methodology.

Nevertheless, cognitive interviewers rarely seem to discuss the proper usage of "confirmatory" or "expansive" strategies, as opposed to "cognitive," nor the proper use of feedback and conversational remarks. If these behaviors are prominent in cognitive interviews, then the potential exists for survey researchers to misunderstand what cognitive interviewing actually accomplishes. Furthermore, the potential for interviewer effects, which could threaten the generalizability of findings, is greatest for behaviors that researchers are not adequately aware of and are not training cognitive interviewers to perform effectively and systematically.

It is useful for cognitive researchers to examine closely what actually happens in these interviews. This study represents a first step in this direction, and that process should continue. To the extent that we understand and can give guidance about what behaviors are optimal and acceptable, we can minimize the risk of problems with this methodology.

References

Beatty, P. (In press). "Review of Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research, edited by Norbert Schwarz and Seymour Sudman." Journal of Official Statistics.

Beatty, P., Schechter, S., and Whitaker, K. (1996). "Evaluating Subjective Health Questions: Cognitive and Methodological Investigations." Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 956-961.

Beatty, P., Willis, G., and Schechter, S. (In press). "Evaluating the Generalizability of Cognitive Interview Findings." Statistical Policy Working Papers, Washington, DC: Statistical Policy Office, U.S. Office of Management and Budget.

Conrad, F., and Blair, J. (1996). "From Impressions to Data: Increasing the Objectivity of Cognitive Interviews." Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 1-9.

DeMaio, T.J. and Rothgeb, J.M. (1996). "Cognitive Interviewing Techniques: In the Lab and in the Field." In Schwarz, N. and Sudman, S. (eds.), Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research. San Francisco: Jossey-Bass, 1996.

Forsyth, B.H. and Lessler, J.T. (1991). "Cognitive Laboratory Methods: A Taxonomy." In Biemer, P., Groves, R., Lyberg, L., Mathiowetz, N., and Sudman, S. (eds.), Measurement Error in Surveys. New York: John Wiley and Sons.

Gerber, E.R. and Wellens, T.R. (1997). "Perspectives on Pretesting: 'Cognition' in the Cognitive Interview." Bulletin de Methodologie Sociologique, 55, 18-39.

Nisbett, R.E. and Wilson, T.D. (1977). "Telling More Than We Can Know: Verbal Reports on Mental Processes." Psychological Review, 84, 231-259.

Schechter, S., Beatty, P. and Willis, G. (1997, February). "Cognitive Issues in Subjective Health Reports." Paper presented at the Conference on Cognition, Aging, and Survey Measurement, held in Ann Arbor, MI.

Willis, G. (1994). "Cognitive Interviewing and Questionnaire Design: A Training Manual." Cognitive Methods Staff Working Paper Series, No. 7. Hyattsville, MD: National Center for Health Statistics.