

Designing Questions for a Multi-Mode Survey of Health Plan Members

Vickie L. Stringfellow and Floyd J. Fowler, Jr., Center for Survey Research
Vickie L. Stringfellow, CSR, University of Massachusetts Boston, 100 Morrissey Boulevard, Boston, MA 02125

Keywords: Mode Effects, Multi-mode Data Collection, Question Design

Introduction

The quality and usefulness of survey data are directly related to response rates. In order to achieve high response rates, and therefore survey responses that are representative of the sampled population, it is often desirable to employ a dual-mode data collection protocol. This strategy requires the development of instruments that produce comparable data across modes. We had the opportunity to compare mail and telephone results from two versions of an instrument to test for the effects of mode on survey answers.

The Consumer Assessment of Health Plans Study (CAHPS[®]) was created to gather plan members' reports and ratings of their medical health care. Extensive testing has been done to reduce the effects of mode of administration. As part of this testing, alternative wording was developed for several questions, and some items were split into two questions. This latter technique often makes questions shorter and easier to understand, but may create the potential for mode effects because respondents to the mail version are able to see the follow-up items and can change their previous answers accordingly. This opportunity obviously does not exist for telephone respondents.

Data were collected using a sample of 2000 health plan members in Washington state. The sample was randomly divided among four data collection protocols, which varied in terms of the questionnaire version they were assigned to complete (standard or alternative) and the initial mode of administration (mail or phone). This design allows for the comparison of responses to 1- and 2-question versions between and within modes of administration.

Methods

Survey Instruments

The CAHPS 2.0 instrument was administered to half the sample, and will be referred to here as the "standard" version.

The "alternative" version was designed to test wording variations of several items and response tasks, and to test the ordering of certain sections. One of the tests involved two questions in the standard instrument that are formatted to ask how much of a problem, if any, the respondent has had with various aspects of getting health care. The response task "big problem", "small problem", "not a problem" is used for these items. In cognitive testing and focus groups, it was learned that some people who have had difficulty getting care or services will not report this as a problem if the issue was eventually resolved. In an effort to reduce this tendency, a variation of this format was tested. Each "problem" question was broken into two questions: The first uses a yes/no response task and asks if the respondent had difficulty with some aspect of care. The follow-up question asks if this was a big problem or a small problem.

The effectiveness of this dual-question variation is important for another reason. In the Spanish version of CAHPS, all items using the "problem" response task were translated into two questions that mirror the test questions

mentioned above. The comparability of English responses to the single item and Spanish responses to the two items is an issue that has not been resolved. Based on the results of this experiment, it may be possible to resolve this issue by using the dual-question format in the English version.

The complete text of the standard and alternative questions can be found in Table 1.

Sample

A sample was drawn of 2000 Washington state employees who were covered by a single commercial health plan. Half the sample was randomly selected to complete the standard CAHPS instrument and the other half was assigned to complete the alternative version. Half of each subsample was then randomly assigned to receive the instrument by mail, the other half was assigned to complete a telephone interview. All sampled members were at least 18 years of age. In situations where two or more individuals were selected from the same household, one was randomly chosen to remain in the sample and the others were replaced.

Data Collection

Members assigned to the mail protocol were sent an introductory letter explaining the purpose of the study and notifying them that they would receive a questionnaire within the next few days. This was followed by an initial survey packet which included a cover letter, fact sheet with answers to commonly asked questions, questionnaire (standard or alternative), and business reply envelope. About a week later, everyone was mailed a postcard thanking them for their participation and asking for their cooperation if they had not yet completed the questionnaire. Three weeks after that, a second survey packet was mailed to everyone who had not yet responded.

Members assigned to the phone protocol were sent an introductory letter explaining the purpose of the study and notifying them that an interviewer would attempt to contact them soon to conduct a telephone interview. Calls were made to directory assistance for members for whom no phone number was available from administrative records and for whom this phone number was incorrect. A minimum of six calls were made on different days and at different times to reach sampled members. The survey was conducted using computer-assisted telephone interviewing (CATI) software.

After the initial field period was over, attempts were made to contact nonrespondents via the mode to which they had not been assigned. Questionnaires were mailed to telephone nonrespondents, and phone calls were made to mail nonrespondents.

Results

Response Rates

The results and response rates for both modes of data collection are presented in Table 2. As noted, sampled members who could not be located are included as eligible cases in the response rates. The last row of Table 2 presents the overall response rates including completed interviews from both modes of data collection, while the penultimate row presents the response rates for the initial mode of data collection. It can be seen that all response rates are similar

within modes. For this paper, only data from respondents to the initial mode were analyzed.

Differences Within Mode

Comparisons of responses to the single question and dual-question formats within each mode are presented in Table 3. It can be seen that the different versions yielded nearly identical results from respondents to the mail protocol. For phone respondents, however, there were significant differences for both items. In both cases, respondents to the single question format were more likely to report a problem than were respondents to the dual-question format.

Differences Between Modes

The results in Table 3 were tested for differences between modes. Mail and phone respondents to the standard questionnaire were compared, and no significant mode differences were found. Mail and phone respondents to the alternative instrument were also compared, and again there were no significant differences between respondents to the two modes.

Overall Differences Between Formats

To test the overall effect of the standard versus alternative question forms, respondents to both modes were combined. The results presented in Table 4 indicate that for both items, more people reported having a problem with the single question than with the dual-question format. The results for both tested items approached but did not reach significant levels.

Conclusion

The impetus for testing the dual-question format for CAHPS items using the “problem” response task was twofold. First, it was thought that respondents who had some difficulty getting health care would be more likely to report this as a problem with the alternative wording. In fact, the opposite finding emerged: Fewer problems overall were reported with the dual-question format than with the single question, though the differences were not significant. Perhaps more importantly, significant mode effects were found with this format. Telephone respondents were less likely to report problems than respondents to the self-administered questionnaire. This is possibly due to the fact that mail respondents are able to read ahead and change their answers to the screening question after seeing the follow-up items.

The second purpose for this test was to inform the decision about whether to change the English questions to mirror the format of the Spanish translation, which uses two questions for every item that employs the “problem” response task. Asking two questions in Spanish and a single question in English is likely to affect the comparability of answers. Again, the effect of mode on the dual-question format is not only a strong argument against changing the English instrument, but also raises concerns about the existing Spanish translation.

Based on these data, the dual-question format is not a viable option for items using the “problem” response options. Further, the results here indicate that it would be prudent to reexamine the Spanish versions of these items and attempt to find a valid translation that retains the single question format.

Table 1. Text of Standard (Single-question) and Alternative (Dual-question) Formats

Domain	Standard	Alternative
Problem getting necessary care	In the last 12 months, how much of a problem, if any, was it to get the care you or a doctor believed necessary? A big problem A small problem Not a problem	In the last 12 months, was there <u>any difficulty</u> getting care, tests, or treatment you or a doctor believed were necessary? Yes No → If No, Go to X Was <u>difficulty</u> getting the care, tests, or treatment you needed <u>a big problem</u> , <u>a small problem</u> , or <u>no problem at all</u> ? A big problem A small problem No problem at all
Problem with delays while waiting for plan approval	In the last 12 months, how much of a problem, if any, were delays in health care while you waited for approval from your health plan? A big problem A small problem Not a problem	In the last 12 months, were there any <u>delays</u> in health care while you waited for approval from your health plan? Yes No → If No, Go to X Were delays while you waited for approval <u>a big problem</u> , <u>a small problem</u> , or <u>no problem at all</u> ? A big problem A small problem No problem at all

Table 2. Results and Response Rates by Initial Mode of Contact and Instrument Version

	Mail		Phone	
	Standard	Alternative	Standard	Alternative
Original sample	500	500	500	500
Ineligible	6	5	5	9
Completed by mail	278 ³	279 ³	51	43
Completed by phone	57	68	263 ³	268 ³
Refusal	48	54	69	67
Not able to locate	57	44	70	68
Other non-interview ¹	6	6	11	8
Field limit	48	44	31	36
Response rate for initial mode ²	56%	56%	53%	55%
Overall response rate ²	68%	70%	63%	63%

¹ Includes 5 people who were unable to complete the survey in English, 14 people who were away for the duration of the study, and 12 people who were too ill to participate.

² (Completed by mail + Completed by phone)/Original sample - Ineligible

³ Respondents included in analyses for this paper.

Table 3. Single-question vs. Dual-question Responses Within Mode of Administration

		Problem getting care believed necessary by doctor			Problem with delays while waiting for plan approval		
		<i>n</i>	Single Question	Two Questions	<i>n</i>	Single Question	Two Questions
Mail	Reported a problem	54	11%	10%	41	8%	8%
	No problem	469	89%	90%	481	92%	92%
	Total	523	100%	100%	522	100%	100%
Phone*	Reported a problem	54	14%	8%	33	9%	4%
	No problem	434	86%	92%	454	91%	96%
	Total	488	100%	100%	487	100%	100%

* Respondents to single- and dual-question instruments significantly different at the $p > .05$ level

Table 4. Single-question vs. Dual-question Responses Combined Across Modes of Administration

	Problem getting care believed necessary by doctor			Problem with delays while waiting for plan approval		
	<i>n</i>	Single Question	Two Questions	<i>n</i>	Single Question	Two Questions
Reported a problem	108	13	9	74	9	6
No problem	903	87	91	935	91	94
Total	1011	100%	100%	1009	100%	100%