

EVALUATING MODE EFFECTS ON A SURVEY OF BEHAVIORAL HEALTH CARE USERS

Vickie L. Stringfellow, Floyd J. Fowler, Jr., Brian R. Clarridge, Center for Survey Research
Vickie L. Stringfellow, CSR, University of Massachusetts at Boston, 100 Morrissey Blvd, Boston, MA 02125

Key words: Social Desirability, Mode Effects, Behavioral Health

One underlying issue for telephone surveys, particularly those pertaining to sensitive issues, is that social desirability will influence responses. The effects of social desirability on answers may be expected to be greater for a study of behavioral health care than for a study of general health care. We had a unique opportunity to test the effects of telephone administration versus self-administration for similar items in two instruments, one pertaining to medical health care and the other to behavioral health care.

The Consumer Assessment of Health Plans Survey (CAHPS) instrument was created to gather plan members' reports and ratings of medical health care. Extensive testing and design has been done to reduce the effects of mode of administration. In a study of health plan members in Washington state, a probability sample of 1000 commercial plan members was selected to complete the CAHPS instrument. Half the sample was randomly assigned to receive the instrument by mail, the other half was assigned to complete a telephone interview.

The Experience of Care and Health Outcomes (ECHO) instrument was designed, using many questions parallel to CAHPS, to obtain the same type of information about the behavioral health care plan members may have received. Data were collected using the ECHO instrument in Massachusetts, where 1000 commercial plan members were selected to complete the interview by telephone. An additional sample of 1500 commercial members was mailed a self-administered questionnaire.

Using data collected from both studies, we are able to compare the effects of mode on the parallel CAHPS and ECHO items and answer several research questions: Does mode of administration have an effect on answers? If so, do these effects indicate that telephone respondents to the behavioral health survey may be motivated by social desirability?

METHODS

Sample

For the ECHO sample, members of two commercial health plans in Massachusetts were randomly assigned to either mail (n = 1500) or telephone (n = 1000) data collection protocols. Only members who had received

behavioral health services between January 1, 2000 and June 30, 2000 were eligible for selection. For the CAHPS sample, Washington state employees who were covered by a single commercial health plan were randomly assigned to either mail (n = 500) or telephone (n = 500) data collection protocols. All selected members for both the ECHO and CAHPS samples were at least 18 years of age. In situations where two or more individuals were selected from a single household, one was randomly chosen to remain in the sample and the others were replaced.

Survey Instruments

Excluding demographics and administrative items, the CAHPS instrument is comprised of 38 questions about health care received in the last 12 months. The version of the ECHO instrument used for this study included 74 questions about mental health counseling or substance abuse treatment received in the last 12 months. Because the CAHPS instrument was used as a starting point for the creation of the ECHO, there are many questions common to both. Twenty-six items designed to gather reports and ratings of care are included in both instruments, and this paper focuses solely on the comparisons of those items.

Data Collection

Members assigned to the mail protocols were sent an initial survey packet consisting of an introductory letter, a fact sheet with answers to frequently asked questions about the study, a questionnaire, and a postage-paid business reply envelope. About a week later, everyone was sent a postcard thanking them for their participation and asking for cooperation if they had not yet returned the completed questionnaire. Three weeks after that, nonrespondents were sent a second full survey packet.

Members assigned to the telephone protocols were mailed an introductory letter which explained the purpose of the study and notified them that they would be contacted soon to complete a telephone interview. For members for whom no telephone number was available from administrative records or for whom this number was incorrect, interviewers called directory assistance to search for a current number. Professional interviewers subsequently made at least six attempts to contact each individual for whom a telephone number could be obtained. Calls were made on different days at different times to contact respondents who were difficult to reach. The survey was conducted using computer-

assisted telephone interviewing (CATI) software.

For the CAHPS sample, data collection for both protocols began on December 1, 1999 and ended March 6, 2000. For the ECHO sample, data collection for both protocols began on November 1, 2000 and ended December 31, 2000.

RESULTS

Response Rates

Data collection results and response rates are presented in Table 1. Calculations include all sampled individuals as eligible, including those with inadequate contact information, unless it was specifically learned that a person did not belong to the plan from which they were sampled. The overall response rates for each mode in the ECHO sample differ markedly (20% for mail, 47% for phone). However, the total number of returns is sufficient to permit reliable comparisons of the two groups.

Data Cleaning

Data collected via the mail protocol for each instrument were merged with the data from the computer-assisted telephone interviews. These files were then checked for internal consistency and invalid codes. Separate codes were used for items for which no valid response was obtained ('Not Ascertained') and items which were legitimately skipped based on answers to prior screening questions ('Inapplicable').

Overall Differences by Mode

Mail and telephone responses to all 26 items were compared separately for the CAHPS and ECHO data sets. The Chi-square test was used to identify significant differences between modes for 24 of these items. For 2 items that use a 0-10 rating scale, the mean scores from each mode were compared using an independent samples t-test. Statistically significant differences by mode were obtained for 4 items in the CAHPS results and 6 items in the ECHO results. Only 1 item was found to be statistically significant by mode for both samples.

Direction of Differences

To assess the direction of these differences, reports and ratings that utilized non-dichotomous response options were recoded into dichotomous variables based on the response task. For example, items with the response choices "Never, Sometimes, Usually, Always" were divided into the categories "Always, Not Always". These new variables were then retested for differences by mode using the Chi-square test. Mode effects were again evident in all but 1 of the 4 items from CAHPS and 1 of the 6 items from ECHO. The results of these comparisons are displayed in Table 2.

Trends in the direction of the mode effects are

evident. In every case, for both survey instruments, phone respondents gave more positive answers than mail respondents.

Health Status and Mental Health Status vs. Mode Effects

It is possible that the differences by mode of data collection are the effect of less healthy individuals and those who use more services electing to respond by mail rather than an effect of mode itself on responses. To see if mode continued to have a significant effect on responses after controlling for health status in the CAHPS sample and mental health status in the ECHO sample, we estimated regression models. Linear regression was used for the 7 items that utilize response tasks with more than two options, and logistic regression was used for the 2 items that use the 'yes/no' response task.

For the CAHPS sample, a self-reported rating of overall health which used a 5-category response task (excellent, very good, good, fair, poor) was used to control for health status. The number of visits to doctors' offices or clinics in the last year as reported in the survey instrument was used to measure utilization of services.

For the ECHO sample, mental health status was controlled by using a self-reported rating of overall mental health which used a 5-category response task (excellent, very good, good, fair, poor). The measure of utilization was the number of outpatient visits for counseling or treatment in the last year as reported in administrative data.

The results of the regression analyses for the effects of mode are presented in Tables 3. Overall, after controlling for health or mental health status and utilization of services, mode continued to have a significant effect on only 3 items.

CONCLUSION

Because respondents to a mail data collection protocol self-select, that is they must actively choose whether or not to fill out and return the questionnaire, there is the potential for respondents to be systematically different from phone respondents in ways that are related to the survey content. Previous studies of consumer health care assessments have shown that people are more likely to respond by mail if they have chronic health conditions or use more services than average. Telephone nonresponse seems often to be less related to survey content, as less respondent effort is required and interviewers can often enlist cooperation from less interested respondents. As demonstrated here, the effects of mode of administration in the CAHPS and ECHO surveys are primarily attributable to differences

in who responds to mail and phone requests rather than to the effects of the mode on the way people answer questions. As such, social desirability does not appear to be a factor for telephone responses to the ECHO instrument. The rate of significant differences due to mode of administration was equivalent for the ECHO and CAHPS instruments, both before and after controlling for health status.

The difference in response rates between the CAHPS and ECHO samples may be an indicator of potential respondents' reluctance to participate in a behavioral health survey, which they may perceive to be a more sensitive topic. Once they decide to participate, however, their responses do not appear to be influenced by mode or social desirability to a greater extent than responses to the CAHPS general health care survey. While the topic of behavioral health care may be a sensitive one, the questions included in the instrument are not. The fact that ECHO survey items gather reports and ratings of the quality of care received, rather than personal or diagnostic information, likely makes the responses less susceptible to social desirability effects in general.

Table 1. Final Results and Response Rates by Sample and Mode

	Initial Sample	Ineligible	Not Able to Locate	Refusals	Field Limit	Other Results	Completed Interviews	Response Rate
CAHPS								
Mail	500	4	20	18	178	3	277	56%
Phone	500	6	70	62	89	13	260	53%
Total	1000	10	90	80	267	16 ^a	537	54%
ECHO								
Mail	1500	–	50	3	1153	1	293	20%
Phone	1000	4	151	240	128	10	467	47%
Total	2500	4	201	243	1281	11 ^b	760	30%

^a Includes 4 members who were away for the duration of the field period, 6 members who were considered too ill or cognitively unable to participate in the study, 2 members who did not speak English well enough to participate, and 4 members who did not complete at least 80% of the instrument.

^b Includes 3 members who were away for the duration of the field period, 3 members who were considered too ill or cognitively unable to participate in the study, 1 member who did not speak English well enough to participate, and 4 members who did not complete at least 80% of the instrument.

Table 2. Dichotomous Variations of Items with Significant* Differences by Mode

	CAHPS		ECHO	
	Mail %	Phone %	Mail %	Phone %
Needed care right away	44 (277)	34 (259)	37 (272)	29 (427)
Never waited 15+ minutes past appointment time	30 (246)	49 (222)	–	–
Doctors always listened carefully**	61 (246)	67 (224)	–	–
Looked for written health plan information	59 (275)	50 (260)	–	–
Always got an appointment as soon as wanted	–	–	45 (255)	54 (416)
Doctors always showed respect**	–	–	77 (257)	81 (411)
Had a problem understanding written plan information	–	–	46 (68)	25 (119)
	Mean	Mean	Mean	Mean
Mean rating (0-10) of all care	–	–	7.89 (257)	8.15 (411)
Mean rating (0-10) of plan	–	–	7.13 (268)	7.78 (423)

* p<.05

** Significance achieved only with original response options, not with dichotomized variation

Table 3. Significance of Mode After Controlling for Health Status and Utilization of Services in Regression Analyses of Items with Significant* Differences by Mode

	<u>CAHPS</u>	<u>ECHO</u>
Linear Regression		
p (Standardized Beta)		
How often waited 15+ minutes past appointment time	.006* (-.130)	–
How often doctors listened carefully	.141 (.069)	–
How often able to get appointment as soon as wanted	–	.731 (.013)
How often doctors showed respect	–	.098 (.064)
Problem with plan information in written materials	–	.003* (.215)
Rating (0-10) of all care	–	.428 (.030)
Rating (0-10) of plan	–	.001* (.125)
Logistic Regression		
p (Beta)		
Needed care right away	.102 (-.307)	.179 (.231)
Looked for written health plan information	.163 (-.252)	–

* Statistically significant effect at the $p < .05$ level