

Effects on Level and Quality of Response of the Inclusion of "Don't Know" Boxes in Factual Questions in a Mail Questionnaire

Gail S. Poe, Isadore Seeman, Joseph McLaughlin, Eric Mehl, and Michael Dietz

In the development of each question in a survey questionnaire the survey designer must decide, in addition to question wording, what response categories are most appropriate. Usually closed response categories are preferred to facilitate coding and analysis. The survey designer's choice of specific response categories can be as critical as choosing the specific question wording. If inappropriate or inadequate response categories are used, the reported data can severely misrepresent reality.

In selecting the specific response options the survey designer may choose to either include or exclude a "don't know" (DK) box. (Alternatively, a screener question may be used that permits the respondent to state that he/she has no opinion or no access to information.) In opinion-type questions, it is usually appropriate to allow a DK option because some people truly have never thought about an issue or have never formulated a position or point of view.

Previous research on effects of a DK response option has focused on attitude questions, administered by interviewers, involving the use of "filter" questions to allow respondents to state that they have no opinion (or have not thought about an issue, etc.). (Bishop et al., 1986, 1983, 1980, 1978 and Schuman and Presser, 1981, 1978)

An issue which has not been extensively investigated is the effect of the inclusion of a DK option for factual questions. Respondents may legitimately not know an answer because they never knew the information or they can no longer remember the information. Including DK boxes for a factual question might result in more accurate response distributions. It is preferable to have respondents state that they don't know an answer to a question rather than making a wild guess or leaving the item blank. On the other hand, including DK boxes might result in higher unknown rates, including both blanks and stated DK's. Respondents may be more likely to take the easy way out rather than taking time to search memory, to refer to records, or to ask someone else who may possess the needed information, particularly with self-administered questionnaires.

This article presents the findings of an experiment to measure the effects of the inclusion of DK boxes on the level of response to factual questions in a mail survey.

STUDY DESIGN

The experiment was included in a large pretest conducted by the National Center for Health Statistics (NCHS) in 1985 for the 1986 National Mortality Followback Survey. A sample of 1,360 death certificates of persons 25 years of age and older was selected from four States: Illinois, Vermont, Virginia, and New Mexico. The respondents who provided information about the decedents were usually close relatives. The sampling frame was death certificates filed in September and October of 1984 as part of the Current Mortality Sample (CMS) in the National Vital Registration System. The CMS is a 10 percent systematic sample of deaths which is selected on a state-by-state basis and tabulated and analyzed by NCHS. In selecting the sample for the pretest, blacks were oversampled at a rate of 1.8 times the rate for

persons other than black; and all deaths from ischemic heart disease at ages 25-44 were selected.

The pretest sample was not designed to be a probability sample of all adult deaths in the U.S. as the main survey sample is. However, a comparison of the total pretest sample (weighted to adjust for differential sampling rates) with all deaths in the U.S. for 1983 shows that the proportion of all deaths in the sample did not differ significantly from U.S. deaths with regard to sex or age. A difference did occur for race, with blacks 3.4 percent higher in the sample than in the U.S. This was a result of the selection of Virginia and Illinois, which have higher than average black populations.

The major topic areas in the survey are socioeconomic status, health care in the last year of life, and health practices. Because death certificates are primarily legal documents, they contain only very limited information. The National Mortality Followback Survey is intended to obtain a great deal of additional information about decedents for use by health planners and policymakers. The major sponsors of the Survey, besides NCHS, include the Health Care Financing Administration and several other agencies and organizations in the Public Health Service. The Bureau of the Census was the data collection agency designated by NCHS for both the pretest and the main survey.

To obtain an eligible respondent the informant on the death certificate was first sought out. The informant was usually a close relative of the decedent. If the name of an informant was not given on the death certificate, funeral directors, attending physicians, coroners, and others were called to obtain the name and address of a close relative or other suitable respondent. The last resort was to address the questionnaire to "next-of-kin" of the named decedent at the decedent's address listed on the death certificate.

A self-administered questionnaire was mailed to the informant by first class government-paid mail about six months after the death. A postage-paid return envelope was included. Ten days later a letter was mailed to all subjects thanking them if they had returned the questionnaire and reminding them to do so if they had not. One month after the initial mailing, a second copy of the questionnaire was mailed to nonrespondents. (Additional followups were made by personal and telephone interviews but those results are not included in this paper.)

To study the effects of using DK boxes, the sample was randomly split, with one-half of the respondents receiving a questionnaire with DK boxes and the other one-half receiving a questionnaire without DK boxes. Respondents with DK boxes were told in printed instructions at the beginning of the questionnaire that if they didn't know the answer to a question they were to mark the DK box if one was provided, or to put a question mark (?) in the answer space. Respondents without DK boxes were instructed to put a question mark in the answer space if they didn't know an answer. Any indication by the respondent that an answer was not known to them was coded as a DK.

As part of the pretest, a large subsample (22 percent) of the completed interviews were reinterviewed and reasked a subset of the original

questions. The original respondent was contacted by a senior interviewer or an interviewer supervisor within one or two weeks of the original interview. The reinterviews, all completed by telephone, were conducted in the same manner regardless of the version of the questionnaire used initially. Although DK boxes were printed in the reinterview questionnaires, when the reinterviewer read the question she did not state the response option of don't know. Most of the items selected to be included in the reinterview were chosen because of concern of the survey planners over the respondents' ability to report the information. A very valuable function of the reinterview program is that it provides a measure of the reliability of the survey responses. By comparing the original interview responses to the responses to the same questions asked in the reinterview, the simple response variance contribution to the total survey error may be calculated (Koons, 1973). For each item included in the reinterview the simple response variance, also called the gross error rate, can be compared for the questionnaire version with DK boxes versus the version without. Lower gross error rates for one version versus the other would be evidence of superior quality. For all comparisons made in this paper a significance level of .05 was used.

RESULTS

The total mail response rate was about the same for the version with DK boxes (61.5 percent) as the version without (58.2 percent) (Table 1). The fail-edit rate, one measure of the quality of responses, was also similar for both formats. (A manual edit was performed upon receipt of the questionnaires by examining 18 "key" items. If four or more of these 18 items were blank, illegible, or had inconsistent information the form was classified as a "fail-edit" and a telephone call was made to the respondent to obtain more complete responses.)

Responses for each question were classified as either an appropriate skip, a substantive (or useful) response, a DK, a blank (when an entry was required), or other nonsubstantive response. (The other category included illegible entries, multiple responses, etc.) For each questionnaire returned, the percents of substantive replies, DK's, and blanks were computed and these respective percents were averaged across all mail returns. For each questionnaire version, the sum of the average percents of DK's, blanks, and substantive replies was 99.9 percent. (The "other" category accounted for only 0.1 percent of each version.) The average percent of items marked either DK or with a question mark was significantly higher for the version with DK boxes (7.1 percent) than for the version without (1.9 percent). Since the sum of the average percents had to total 100 percent for each version, this difference in the percent of DK replies between the versions had to be balanced by differences between the versions in either the average percent of blanks, the average percent of substantive replies, or both. The average percent of items left blank was a little higher but not significantly so for the questionnaire without DK boxes (13.3 percent versus 11.3 percent).¹ The version without DK boxes had an average percent of substantive responses 3.2 percent higher than the version with DK boxes; however, this difference is not statistically significant.

The following variables were examined to determine if they interacted with inclusion or exclusion of DK boxes and no interactions were found: decedent's age, race, Hispanic origin, sex, marital status, and cause of death, and relationship of respondent to decedent and number of years respondent lived with the decedent.²

Although the average percent of items with substantive responses was not significantly higher for the version without DK boxes, for many specific items the percent of substantive responses for this version was appreciably higher. For some very important items the percent of substantive replies in the version with DK boxes was so low as to present considerable concern to analysts. For example, in the question on number of cigarettes smoked the version with DK boxes produced only 69.7 percent substantive replies (compared to 87.7 percent for the version without DK boxes). For one-fourth of the items the percent of substantive replies was significantly higher for the version without DK boxes (Table 2). For four items in the questionnaire the percent of substantive replies was 25 or more percent higher for the version without DK boxes. For 51 items (27.2 percent) the percent of substantive replies was at least 5 percent higher and for 40 of these 51 items the difference was statistically significant. For only one item was the percent of substantive replies significantly higher for the format with DK boxes. There was no single characteristic (eg. potentially sensitive) that would typify the questions which had higher substantive responses in the format without DK boxes.

In order to investigate whether the inclusion or exclusion of DK boxes had an effect on the distribution of substantive responses, for each questionnaire item the distribution of substantive responses was compared between the version with DK boxes and the version without. Significant differences between the versions in substantive response distributions would be evidence of response bias in one or both versions. In most of the items (89 percent) there were no significant differences in substantive responses in any response categories. (See example in Figure 1.) (Forty-one items were excluded from this analysis because there were 20 or fewer substantive responses per version, they were open ended, or the items allowed multiple responses.) In 17 questionnaire items (11 percent) there were significant differences in substantive responses in one or more response categories. (Table 3) Three of these 17 items had 30 percent or more blank or DK entries and were dropped from the main survey. For half of the remaining 14 items the differences in percents for specific substantive replies in the version with DK boxes versus none were less than 10 percent. Hence, there were few large differences in substantive response distributions between the two questionnaire versions.

No patterns of differential responses were detected in the comparison of substantive response distributions between the 2 versions (for example, a higher proportion of "yes" responses or longer time intervals in one version or the other).

For each item included in the reinterview a gross error rate was calculated for both questionnaire versions. The gross error rate is the number of subjects in which the reinterview response disagreed with the original response divided by the total subjects. The average gross error rate across all items for the format with the DK boxes was about the same (15 percent) as the version without (13 percent). None of the items included in the reinterview were ones that had significant differences in distributions of substantive responses as shown in Table 3.

CONCLUSION

Based upon our analysis, the self-administered questionnaire without DK boxes is preferred for the following reasons: (1) an appreciably higher rate of usable or substantive responses for many items; (2) less

imputation for missing responses would be required, thus reducing the potential for introducing error into the data set; (3) the use of DK boxes results in few large substantive changes in response distributions; (4) there is no appreciable difference in the response error rate if DK boxes are excluded; (5) the absence of DK boxes did not affect the overall questionnaire return rate; and (6) without the DK boxes the questionnaire is less cluttered, slightly shorter, and the skip instructions are easier to follow.

It would be worthwhile to repeat this experiment using other survey populations because of the uniqueness of our population (relatives of recently deceased persons). It would be especially interesting to learn if the present findings would be observed in a study of self rather than proxy respondents.

References

- Bishop, George F., Alfred J. Tuchfarber, and Robert W. Oldendick: "Opinions on fictitious issues: The pressure to answer survey questions." Public Opinion Quarterly 50: 1986. Pp. 240-250.
 - Bishop, George F., Robert W. Oldendick, and Alfred Tuchfarber: "Effects of filter questions in public opinion surveys." Public Opinion Quarterly 47: 1983. Pp. 528-546.
 - Bishop, George F., Robert W. Oldendick, and Alfred Tuchfarber, and Steven E. Bennett: "Pseudo-opinions on public affairs." Public Opinion Quarterly 44: 1980. Pp. 198-209.

- Bishop, George F., Robert W. Oldendick, and Alfred Tuchfarber: "Effects of question wording and formation on political attitude consistency." Public Opinion Quarterly 42: 1978. Pp. 81-92.
 - Koons, D.A.: "Quality control and measurement of nonsampling error in the health interview survey." Vital and Health Statistics, Series 2, No. 54, Washington, D.C., U.S. GPO. 1973.
 - Schuman, Howard, and Stanley Presser: Questions and Answers in Attitude Surveys. New York: Academic Press. 1981.
 - Schuman, Howard, and Stanley Presser: "The Assessment of 'No Opinion' in Attitude Surveys." Pp. 241-75 in K.R. Schuessler (ed.), Sociological Methodology 1979. San Francisco: Josey Bass. 1978.

Text Footnotes

1. The relatively high rate of blanks in this study is primarily due to two reasons: a) the method of calculating the percent blank and b) question formatting. The proportion of blanks was calculated as the number of blanks divided by the number of potential responses. For questions with previous screener questions, blanks and don't know answers to the previous screener questions were included as potential responses. In addition, many questions appeared to have had an increased blank rate because a matrix type format was used which may have been difficult for some respondents. All matrix type formatting was eliminated for the main survey.
2. Data may be obtained from authors.

FIGURE 1

EXAMPLE OF COMPARISON OF DISTRIBUTION OF SUBSTANTIVE RESPONSES BETWEEN THE VERSION OF THE QUESTIONNAIRE WITH DON'T KNOW BOXES AND THE VERSION WITHOUT^{1/2/}

Question B27:

- During the last year of life, how often did the person have trouble remembering what year it was? Mark (X) only one box.
- All or most of the time
 - Some of the time
 - Never or hardly ever
 - Only in the last few hours or days before death
 - Don't know)

SUBSTANTIVE RESPONSE CATEGORIES	TOTAL	WITH DON'T KNOW BOXES	WITHOUT DON'T KNOW BOXES
Total substantive responses	721 (100.0%)	351 (100.0%)	370 (100.0%)
All/most of the time	67 (9.3%)	28 (8.0%)	39 (10.5%)
Sometimes	78 (10.8%)	37 (10.5%)	41 (11.1%)
Never/hardly	427 (59.2%)	214 (61.0%)	213 (57.6%)
Only last days	149 (20.7%)	72 (20.5%)	77 (20.8%)

^{1/}Mail returns only

^{2/}Unweighted

Table 1-Results of Mailing by Inclusion of "Don't know" Box

Nature of Response	With Don't Know Box	Without Don't Know Box	With Don't Know Box	Without Don't Know Box
	Number		Weighted Percent	
Total Mailed	678	682	100.0	100.0
Total Returned	439	437	66.5	65.3
Filled questionnaires	405	390	61.5	58.2
Pass edit	387	371	59.0	55.6
Failed edit	18	19	2.5	2.6
Refused	14	17	2.1	2.7
Other ^{1/}	20	30	2.9	4.4
No Response	239	245	33.5	34.7

^{1/}Includes new informant identified, postmaster returns, etc.

Table 2. Distribution of Items According to Percent Difference in Rate of Substantive Responses Between Version With and Without DK Boxes

Percent Difference	Number and Percent of Items ^{1/}	
	Total	Significant Differences ^{2/}
25.0 or greater	4 (2.1)	4 (2.1)
15.0 up to 25.0	14 (7.5)	9 (4.8)
10.0 up to 15.0	12 (6.4)	12 (6.4)
5.0 up to 10.0	21 (11.2)	15 (8.0)
0.0 up to 5.0	66 (35.3)	5 (2.7)
Less than 0.0 up to -5.0	53 (28.3)	0 (0.0)
-5.0 up to -15.0	17 (9.1)	1 (0.5)
All differences	187 (100.0)	46 (24.6)

^{1/}Excludes all parts of item B12 because questionnaire design precludes determining appropriate base

^{2/}Significance level of .05

Table 3. Items in Which There Were Significant Differences in Distribution of Substantive Responses

Item	Response Category in Which There was Significant Difference	Version with "Don't Know" Boxes	Version without "Don't Know" Boxes	
		<u>Percent^{1/} of Answers in Specific Response Category</u>		
B20	Duration needed help bathing ^{2/}	2 up to 3 months	2.7	9.8
B20	Duration needed help dressing ^{2/}	2 up to 3 months	4.2	12.2
B30(3)	Problem getting help to care at home ^{3/}	Not a problem	84.8	91.7
		Serious problem	11.5	4.4
B30(4)	Problem paying bills ^{3/}	Not a problem	79.4	86.7
C19	Interval between Alzheimer's diagnosis and death ^{2/}	1 week up to 1 month	10.0	36.2
C21(1)	Interval between diabetes diagnosis and death ^{2/}	3 up to 6 months	10.9	28.0
		10 or more years	34.4	19.5
C21(5)	Interval between chronic lung disease diagnosis and death ^{2/}	1 up to 2 months	12.9	0.0
C23	Interval between last breast exam and death ^{4/}	Less than 1 year	67.2	49.3
		10 or more years	0.0	7.0
C24	Breast self exam frequency ^{5/}	At least once per month	33.9	15.6
C31	Immunized for pneumonia	Yes	26.5	15.5
		No	73.5	84.5
C41	How long smoked regularly ^{6/}	8 up to 11 years	1.8	6.7
		20+ years	78.2	66.0
C42	Peak number cigarettes smoked per day ^{7/}	Under 20	23.1	37.1
D17	Family income ^{8/}	Under \$5,000	20.4	30.2
		\$30,000-\$39,999	7.3	3.5
D27	Years married ^{9/}	6 up to 11 years	1.2	7.4
D30	Number living sisters 25+ years ^{10/}	4-6	13.2	19.1
D32	Age of mother at heart attack ^{11/}	65-74 years	50.0	25.8
D33	Assets ^{12/}	\$1-\$4,999	9.4	20.4
		\$100,000-\$249,999	12.6	5.2

^{1/} Unweighted

^{2/} Substantive response categories: Under 1 week, 1 week up to 1 month, 1 up to 2 months, 2 up to 3 months, 3 up to 6 months, 6 months up to 1 year, 1 up to 6 years, 6 up to 10 years, 10+ years

^{3/} Substantive response categories: Not a problem or not applicable, somewhat serious problem, very serious problem

^{4/} Substantive response categories: Less than 1 year, 1 up to 2 years, 2 up to 3 years, 3 up to 4 years, 4 up to 5 years, 5 up to 6 years, 6 up to 10 years, 10+ years, never

^{5/} Substantive response categories: At least once per month, less frequently than once a month, no

^{6/} Substantive response categories: Never smoked regularly, less than 1 year, 1 up to 3 years, 3 up to 8 years, 8 up to 11 years, 11 up to 20 years, 20+ years

^{7/} Substantive response categories: Under 20, 20, 21-39, 40 or more

^{8/} Substantive response categories: Less than \$5,000, \$5,000-\$9,999, \$10,000-\$14,999, \$15,000-\$19,999, \$20,000-\$29,999, \$30,000-\$39,999, \$40,000-\$49,999, \$50,000 or more

^{9/} Substantive response categories: Under 1 year, 1 up to 6 years, 6 up to 11 years, 11 up to 16 years, 16 up to 21 years, 21 up to 31 years, 31 up to 41 years, 41 up to 51 years, 51 or more years

^{10/} Substantive response categories: None, 1, 2-3, 4-6, 7-10, 11+

^{11/} Substantive response categories: Under 25, 25-44, 45-64, 65-74, 75+

^{12/} Substantive response categories: None, \$1-\$4,999, \$5,000-\$9,999, \$10,000-\$24,999, \$25,000-\$49,999, \$50,000-\$99,999, \$100,000-\$249,999, \$250,000-\$499,999, \$500,000+