

THE EFFECTS OF QUESTION WORDING ON DATA QUALITY: SOME EMPIRICAL FINDINGS

Grady D. Bruce, California State University, Fullerton
Robert A. Peterson, The University of Texas at Austin

ABSTRACT

Responses to a question worded subjectively and objectively and included in mail and telephone surveys are examined. Question wording produced no significant differences in responses, while mode of administration did. Several interpretations of these findings are discussed, as well as their implications for marketing research.

INTRODUCTION

Marketing researchers have been concerned for some time with the quality of data obtained from consumer surveys (e.g., see Peterson and Kerin 1981). The wording of questions in such surveys has long been a major suspected source of data-quality problems; but, as Hitlin recently noted, "comparatively little empirical work has been published to document many suspected wording effects" (1976, p. 39). One particular effect-- that caused by subjective versus objective (personal versus impersonal) question wording--was studied in the 1940s and 1950s, but virtually ignored since that time (Blankenship 1940 and Hubbard 1950). Early writers (e.g., Droba 1931) believed that the subjective question form (i.e., inclusion of the word "you" in the question) produced more accurate responses than the impersonal form. A major test of this hypothesis, however, failed to support it: the objective form produced more valid and reliable responses (Blankenship 1940). Another test, however, concluded that the subjective form was more reliable (Hubbard 1950). Such conflicting findings--and the paucity of research on the effects of subjective versus objective wording--stimulated the authors to conduct this investigation. The lack of empirical work on this and other issues in question wording is actually quite surprising, given the likely impact of question wording on the quality of data collected and, ultimately, on marketing decisions made utilizing such data. It is the purpose of this investigation to examine empirically the effects of objective versus subjective question wording on the results achieved in survey research.

METHOD

Using the split-ballot technique, a question worded both subjectively and objectively was included in two proprietary investigations of financial attitudes and behavior conducted in the early 1970s (at the onset of the "energy crisis"). The alternative question wordings were:

- (Subjective) In your opinion, is the energy crisis real or artificial? Is it...
- (Objective) Is the energy crisis real or artificial? Is it...

Both questions were posed in a closed-end (categorical) format employing the following response categories:

- very real
- somewhat real
- somewhat artificial
- very artificial, or are you
- uncertain

The two investigations which served as study vehicles employed different modes of data collection--telephone interviews and a mail survey. This was done so that the effect of the alternative question forms could be evaluated when they were administered (and experienced by respondents) both personally and impersonally. As a result, it was possible to assess the effect of data collection mode on data quality, as well as the effect of the interaction of question form and mode. Accordingly, both forms of the question were included in each of the two investigations within the framework of an experimental design.

Both the mail survey and telephone interviews were conducted using standard research practices. The same population was sampled in both studies (the population of a southwestern SMSA). For the mail survey the sampling frame consisted of a current metropolitan phone directory. Potential questionnaire recipients were selected by means of a systematic probability sampling technique, the equivalent of simple random sampling given the alphabetical ordering of the directory. A sample of 3,840 respondents was selected. For the telephone survey telephone numbers were selected using a variant of random digit dialing. Interviews with 1,170 respondents were utilized in this study.

One important difference between the studies was in their timing. The telephone interviews were conducted at three different times with three-month intervals separating them. The mail survey was conducted in the middle of this time span. This introduced the potential confounding effect of a real change in attitudes which could mistakenly be attributed to the method of administration as a result of the combining process. To evaluate this risk, the response distributions for the three telephone surveys were checked for stability before the samples were combined. No substantive differences were found, so the samples were combined and compared as one with the result of the mail survey.

Another difference between the two studies--a clear limitation of this investigation--was differences in response rates. The response rate for the telephone interviews averaged 80 percent of the working numbers contacted; for the mail survey, the response rate was 27 percent. It is not possible to evaluate the effect of this difference on the findings presented below.

FINDINGS

Responses were cast in a three-way cross-classification for analysis; frequencies and percentages are reported in Table 1. Rather than assuming interval data

TABLE 1
RESPONSES TO GASOLINE SHORTAGE QUESTION

	Telephone			
	Is		In Your	
	No.	%	No.	%
Very Real	115	20	120	21
Somewhat Real	152	26	160	27
Somewhat Artificial	132	23	135	23
Very Artificial	96	16	112	19
Uncertain	90	15	58	10
TOTALS	585	100%	585	100%

	Mail			
	Is		In Your	
	No.	%	No.	%
Very Real	91	18	92	18
Somewhat Real	117	23	130	26
Somewhat Artificial	153	30	140	28
Very Artificial	112	22	112	22
Uncertain	33	7	29	6
TOTALS	506	100%	503	100%

and utilizing a standard ANOVA approach, the investigators used a chi-square decomposition technique to assess the statistical significance of the difference. Chi-square was computed first for variations due to the two basic sources (question wording and mode of administration) and for the total table; the residual was assigned to the interaction effect. Results of the analysis are reported in Table 2. As can be seen from this table, the difference in data collection modes was statistically significant, while question wording and interaction effects were not.

TABLE 2
RESULTS OF CHI-SQUARE ANALYSIS

Source	X ²	df	P
Question	7.8	4	.10
Collection mode	39.4	4	<.001
Question- Collection Mode	4.2	4	>.30
TOTALS	51.4	12	<.001

The nature of this mode effect can be seen most easily if the effects of question wording are removed and if similar attitudinal positions are combined; this is done in Table 3. As the table shows, a substantially higher percentage of respondents in the mail survey expressed the opinion that the energy crisis was artificial than did those in the telephone interviews.

TABLE 3
COLLAPSED FINDINGS

	Telephone Percent	Mail Percent
Real	46.8	42.6
Artificial	40.6	51.2
Uncertain	12.6	6.1
	100.0	100.0

Substantially lower proportions of responses appeared in the "real" and "uncertain" categories. Past research supports several interpretations of this finding.

First, the finding may be due to the fact that respondents in the mail survey responded anonymously and thus were willing to express less socially desirable opinions (in this case, mistrust of established companies and the free-enterprise system). Wildman recently noted, "Respondents who can be identified may give more 'socially desirable' responses than those who remain anonymous" (1977, p. 75).

Another interpretation, one also based on social desirability, is rooted in the extent to which there is personal contact between researcher and respondent. In this context, the comparison is usually made between self-administered questionnaires and face-to-face interviews. Sudman and Bradburn say, "Because self-administered questionnaires are more private and do not require a direct revelation of self to another person who is physically present, we would expect that they would be less subject to problems of self-presentation than would face-to-face interviews" (1974, p. 40). They go on to position telephone interviews between these other two collection modes on the dimension of privacy and cite a number of studies supporting the general hypothesis that the more personal the interviewing situation, the more respondents exhibit socially desirable (conforming) behavior in their responses. In support of this contention, Nuckols (1964), for example, reported that respondents were much more willing to report unfavorable experiences with insurance agents in a mail survey than in a personal interview. Therefore, under the assumption that the social norms at the time of the present study emphasized trust of our dominant economic and political institutions in general, the findings are consistent with those of previous investigators.

A third interpretation relates more directly to how the "uncertain" category is treated. Up to this point it has been treated as scale dependent--the midpoint of a scale involving difficult discriminations. A response in this category may, however, reflect something other than a difficult discrimination (Coombs and Coombs 1976; Kalton, Collins, and Brook 1978; Schuman and Presser 1981). If the uncertain response is regarded as a "don't know" response, then a host of additional concerns emerges. Bogart states that a "don't know" response to a survey question often means "don't want to know" or "I don't want to get involved" (1967, p. 344). This may be due to the sensitivity of the question and the respondent's unwillingness to reveal how (s)he feels, or to conflicts the question induces in the respondent between incompatible values (Coombs and Coombs 1976). Several prior studies have investigated the relationship between respondent characteristics and mode of data collection on don't know responses. The question raised by the current findings, in light of the various explanations which have been offered for don't know responses, is whether the increased personal involvement in a telephone survey dealing with controversial or sensitive questions will heighten respondent uncertainty relative to a mail survey. Alternatively stated, does the lack of personal contact in a mail interview attenuate respondent uncertainty? This position is in direct contradiction with that taken by Francis and Busch: "mail questionnaires may, qua instrument, evoke a higher NSP (non-substantive response) than do other data-gathering forms" (1975, p. 208).

A final interpretation of the findings--one which also involves the treatment of uncertain responses as independent of the measurement scale--is based on the opportunity which mail surveys afford the respondent to formulate a position where one has not existed prior to the survey. Cahalan cites this--the opportunity to "prepare considered answers"--as an advantage of mail surveys over personal interviews (1951, p. 578). According to this logic, mail surveys may be expected to produce fewer uncertain responses on topical issues (which the "energy crisis" was at the time of data collection), on attitudes about which public opinion has yet to crystallize. The present findings may also be interpreted as supporting this contention.

CONCLUSIONS AND IMPLICATIONS

Prior tests of the effects of objective and subjective question wording on data quality may attribute more power to these wording options than they in fact possess. In this study no statistically significant differences were found between responses to questions worded in the two ways. Mode of administration, however, produced significant (and instructive) differences in survey results. Several interpretations of these differences were discussed. In general, the findings add to the stream of literature which expresses concern over interviewer "presence" and interviewer-interviewee interaction on data quality. And they stand in contrast to studies reporting no differences due to mode of administration (e.g., Hochstim 1967).

The apparent capacity of the mail survey to free respondents to express deviant attitudes or responses which may not be pleasing or agreeable has several implications. First, in studies of consumer satisfaction/dissatisfaction and disconfirmed expectations, mail surveys may produce more accurate responses than either telephone surveys or face-to-face interviews, especially in situations where consumer dissatisfaction is high. Second, and closely related, since respondents tend to be less willing to express negative or socially undesirable attitudes in face-to-face (or telephone) interviews, acquiescence (agreeing-response) bias may be greater in these modes than in mail surveys whenever Likert-type, agree-disagree scales are used.

Third, studies of the consumption rates of products (or activities) that are highly sensitive to social norms and values (for example, consumption of alcoholic beverages or church attendance) may achieve more accurate results when administered through mail surveys. And finally, for products subject to high reference group influence, it seems reasonable to suggest that respondents may be less likely to express approval of radically new product concepts (perceived as deviant ideas) when such concepts are presented in personal or telephone interviews compared with mail surveys.

REFERENCES

- Blankenship, A.B. (1940), "The Influences of the Question Form Upon the Response in a Public Opinion Poll," Psychological Record, 3, 349-422.
- Bogart, L. (1967), "No Opinion, Don't Know, and Maybe No Answer," Public Opinion Quarterly, 31, 331-345.
- Cahalan, D. (1951), "Effectiveness of a Mail Questionnaire Technique in the Army," Public Opinion Quarterly, 15 (Fall), 575-578.
- Coombs, C. H. and L. C. Coombs (1976), "Don't Know: Item Ambiguity or Respondent Uncertainty?" Public Opinion Quarterly, 40 (Winter), 497-514.
- Droba, D. D. (1931), "Statements as Opinion Indicators," Sociology and Social Research, 15, 550-557.
- Francis, J. D. and L. Busch (1975), "What We Now Know About I Don't Know," Public Opinion Quarterly, 39 (Summer), 207-218.
- Hitlin, R. (1976), "On Question Wording and Stability of Response," Social Science Research, 5 (March), 39-41.
- Hochstim, Joseph R. (1967), "A Critical Comparison of Three Strategies of Collecting Data from Households," Journal of the American Statistical Association, 62 (September), 976-989.
- Hubbard, A. W. (1950), "Phrasing Questions," Journal of Marketing, 15 (January), 48-56.
- Kalton, Graham, Martin Collins, and Lindsay Brook (1978), "Experiments in Wording Opinion Questions," Applied Statistics, 27 (June), 149-161.
- Nuckols, R. C. (1964), "Personal Interview Versus Mail Panel Survey," Journal of Marketing Research, 1 (February), 11-16.
- Peterson, Robert A. and Roger A. Kerin (1981), "The Quality of Self-report Data: Review and Synthesis," in Review of Marketing 1981, Ben M. Enis and Kenneth J. Roering, editors, Chicago: American Marketing Association, 5-20.