

Attributes of Questions and Interviewers as Correlates of Interviewing Performance



Stanley Presser; Shanyang Zhao

The Public Opinion Quarterly, Vol. 56, No. 2 (Summer, 1992), 236-240.

Stable URL:

<http://links.jstor.org/sici?sici=0033-362X%28199222%2956%3A2%3C236%3AAOQAIA%3E2.0.CO%3B2-J>

The Public Opinion Quarterly is currently published by Oxford University Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/oup.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

ATTRIBUTES OF QUESTIONS AND INTERVIEWERS AS CORRELATES OF INTERVIEWING PERFORMANCE

STANLEY PRESSER

University of Maryland

SHANYANG ZHAO

University of Michigan

A basic tenet of survey research is that respondents must answer standardized questions. Consequently, interviewers are trained to ask questions verbatim. Yet evidence from monitoring studies shows that interviewers frequently depart from the prescribed wording (Brenner 1982; Cannell, Lawson, and Hausser 1975; Marquis 1971).

A few studies have explored the correlates of this error, such as question features or interviewer characteristics. Cannell and Robison (1971) found open questions were less apt to be misread than closed ones. The same effect was observed by Mathiowetz and Cannell (1980) and Bradburn, Sudman, et al. (1979), but it was very small. By contrast, Oksenberg (1981) reported that open questions were *more* likely to be misread than closed ones.

In their experimental comparison with sensitive items, Bradburn, Sudman, et al. (1979) found interviewers made somewhat fewer reading errors when questions were “short” than when they were lengthened by nonsubstantive introductions and when they contained “standard” language as opposed to colloquial terms. In addition, they reported that older interviewers and those with the longest experience made more errors, but these differences were not statistically significant.

The present study replicates parts of this earlier work and extends it to other interviewer and question characteristics. We reasoned that the less exposure interviewers had to a question the less likely it would be read verbatim. If this is true, then questions that are part of frequently invoked skip patterns would be misread more often.

Two considerations led us to expect that the later a question ap-

We thank Kerri Bevilacqua for excellent research assistance and Nancy Mathiowetz and anonymous referees for comments.

peared in the interview the less apt it would be to be asked verbatim. Time into the interview is likely to be associated with respondents expressing restlessness, which may translate into interviewers hurrying to finish, and with greater interviewer fatigue. Both should contribute to error.

Finally we hypothesized that questions asked as part of a series with identical formats should cause departures from the question text. More than most items, questions with a repetitive format violate conventional conversation norms. Thus there may be a temptation to vary the wording so as not to seem mechanical.

Reading questions verbatim is only one part of interviewing. Is successful question delivery related to other aspects of interviewer performance? Fowler and Mangione (1990) reported that correct question reading was related to correct probing and recording of answers in a face-to-face survey conducted by inexperienced interviewers. We do not have measures of probing and recording, but we have indicators of two other facets of interviewing: refusal rate and efficiency. On the assumption that experience should promote most aspects of "good" interviewing, we hypothesized that reading errors would vary with experience, refusal rates, and efficiency.

Methods

Two trained students monitored 79 interviews conducted by 40 interviewers at the Maryland Survey Research Center on a state-wide random digit dial survey in fall 1989. Questions were coded as no change in wording, minor change, or major change. Minor changes included nondirective phrases preceding the question such as "The next question is" as well as small changes within the question. Major changes involved more significant departures from the intended wording. (The use of computer-assisted administration meant there were no questions asked that should have been skipped nor items skipped that should have been asked.) In 10 interviews monitored by both students there was 92 percent agreement on the code.

Each of the survey's 94 questions was scored on five dimensions. The "Q-Error" is the number of times the item was misread divided by the total number of times it was asked in the 79 monitored interviews. "Length" is the number of words in the question. "Position" is the question's number (1 through 94). "Familiarity" is the proportion of times the question was asked in all 982 interviews (not just those monitored). And "Series" is a dichotomy indicating whether the wording is largely identical to the one immediately preceding it.

Each of the 40 interviewers was assigned a value on four variables.

The "I-Error" is the number of times an interviewer departed from exact question wording divided by the total number of questions the interviewer asked while monitored. "Experience" is the number of months worked at the center at the survey's outset. "Refusals" is the number of initial refusals the interviewer obtained during the survey divided by the sum of initial refusals and completed interviews. "Efficiency" is completed interviews divided by hours the interviewer worked on the survey.

Results

More than 91 percent of the 5,619 monitored questions were read exactly as worded. Very few (0.2 percent) were read with major changes in wording; the most frequent errors (8.2 percent) involved minor alterations.

The Q-Error increased with increases in Length ($r = .21, p < .02$) and was higher on Series items (23 percent compared with 7 percent for other items, $t = 5.77, 92 \text{ df}, p < .001$). These are independent effects, as Length and Series were unrelated.

By contrast, Position and Familiarity were unrelated to Q-Error. Only features of the question itself seemed to affect verbatim reading.¹

Finally, I-Error was unrelated to Experience, Refusals, and Efficiency. Factors that affected performance in asking questions are apparently different from those gained through experience and from those that affected success either in gaining respondent cooperation or in working efficiently. (Experience did show a significant negative association with Refusals, but was unrelated to Efficiency.)

These results prompted us to examine whether performance in asking questions was a stable interviewer trait. This was possible since 28 of the 40 interviewers were monitored more than once. The correlation of I-Error in one interview with I-Error in the second interview (or, in cases where we had monitored more than one other interview, the mean across two or more other interviews) was .39 ($p < .02$). But this association was due entirely to inexperienced interviewers ($r = .53, p = .01$). The association for interviewers who had worked on previous surveys was essentially zero ($r = .01, \text{n.s.}$). Thus among novice interviewers the propensity to make reading errors was fairly

1. In a multiple regression predicting Q-Error, the coefficients for both Length and Series were highly significant ($p < .002$), while neither the Position nor Familiarity coefficients attained significance ($p > .25$). Inspection of the data also revealed that the Length-Error relationship was roughly linear.

constant across interviews, whereas for experienced interviewers it was interview-specific.²

Discussion

The level of verbatim reading reported here is similar to that reported by Mathiowetz and Cannell (1980) and Oksenberg (1981), but much lower than that reported by Bradburn, Sudman, et al. (1979), Brenner (1982), and Marquis (1971). This suggests the possibility that centralized telephone interviewing (the mode used in Mathiowetz and Cannell [1980], Oksenberg [1981], and the present study) produces greater adherence to question wording than does in-person interviewing (the mode used in the remaining studies). It may be that the enhanced supervisory opportunities in centralized facilities, or the differences in the demands of telephone and face-to-face communication, affect interviewer performance, but only an experiment will tell for sure.

Our results indicate that questionnaire designers may be able to exert some control over interviewer performance in asking questions. Textbooks commonly exhort researchers to keep questions short, out of a concern for respondent comprehension. But our findings suggest that brevity helps interviewers do their jobs as well.³

Contrary to hypothesis, asking questions verbatim was not associated with an item's placement in the questionnaire nor with how frequently the item was asked. However, the relatively short length of the survey may have made the test of the former hypothesis a weak one.

Also contrary to hypothesis, interviewer experience, refusal rate, and efficiency did not affect accuracy in asking questions. In interpreting these findings it is important to bear in mind that question-asking errors in this study involved mainly minor wording changes, some of which may not have been due to carelessness but to the interviewer's effort to improve the questionnaire (e.g., by avoiding mechanical repetition on series items). It is possible that reading errors like most of the ones observed here may not affect respondent answers and consequently may not be a source of bias.

2. This may be a clue that respondent characteristics (e.g., age) contribute to interviewers' departing from question wording.

3. Though it needs to be noted that on threatening topics, respondents may report more accurately to longer questions (Bradburn, Sudman, et al. 1979).

References

- Bradburn, Norman, Seymour Sudman, and Associates. 1979. *Improving Interview Method and Questionnaire Design*. San Francisco: Jossey-Bass.
- Brenner, Michael. 1982. "Response Effects of 'Role-Restricted' Characteristics of the Interviewer." In *Response Behaviour in the Survey Interview*, ed. W. Dijkstra and H. van der Zouwen. London: Academic.
- Cannell, Charles, Sally Lawson, and Doris Hausser. 1975. *A Technique for Evaluating Interviewer Performance*. Ann Arbor, MI: Institute for Social Research.
- Cannell, Charles, and Sally Robison. 1971. "Analysis of Individual Questions." In *Working Papers on Survey Research in Poverty Areas*, ed. John Lansing, Stephen Withey, and Arthur Wolfe, pp. 236–91. Ann Arbor, MI: Institute for Social Research.
- Fowler, Floyd, and Thomas Mangione. 1990. *Standardized Survey Interviewing*. Newbury Park, CA: Sage.
- Marquis, Kent. 1971. "Effects of Race, Residence, and Selection of Respondent on the Conduct of the Interview." In *Working Papers on Survey Research in Poverty Areas*, ed. John Lansing, Stephen Withey, and Arthur Wolfe. Ann Arbor, MI: Institute for Social Research.
- Mathiowetz, Nancy, and Charles Cannell. 1980. "Coding Interviewer Behavior as a Method of Evaluating Performance." In *Proceedings of the Section on Survey Research Methods*, pp. 525–28. Washington, DC: American Statistical Association.
- Oksenberg, Lois. 1981. *Analysis of Monitored Telephone Interviews*. Research Report to the Bureau of the Census (JSA 80-23). Ann Arbor, MI: Institute for Social Research.