Journal of Applied Psychology
1975, Vol. 60, No. 2, 252-254

Impact of Questionnaire Length, Follow-Up Methods, and Geographical Location on Response Rate to a Mail Survey

Jagdish N. Sheth
Department of Business Administration, University of Illinois

A. Marvin Roscoe, Jr.
American Telephone & Telegraph Company, New York, New York

The impact of two levels of questionnaire length, four follow-up methods, and eight heterogeneous geographical areas on response rate to a mail survey was experimentally assessed. Questionnaires were mailed to a large random sample of telephone customers. For each geographical location, 264 customers ($n = 296$ for southeast Massachusetts) were drawn and equally divided among the eight experimental conditions. No significant differences in response rate occurred between a short, four-page questionnaire and a long, six-page questionnaire. However, the telephone reminder produced the best results, post reminder the second best results, and telephone interviews both with and without prior alert produced the poorest results. Among the geographical areas, the Fort Worth area had the highest response rate ($p < .05$), and Alabama had the poorest ($p < .05$). There were significant interaction effects, $F(21, 63) = 2.20$, $p < .05$, between follow-up methods and geographical areas, suggesting a market segmentation strategy.

While considerable experimental research exists on the rate of response in mail surveys as a function of a variety of manipulations, there is still very little agreement on the factors that are conducive to optimizing the rate of response (Champion & Sear, 1969; Erdos, 1970; Gullahorn & Gullahorn, 1963; Landy & Bates, 1973; Linsky, 1965; Scott, 1961). Our study is one more experimental effort to measure the main and interaction effects of (a) questionnaire length, (b) follow-up procedures, and (c) geographical differences on the rate of return of questionnaires. Specifically, the study was designed to provide answers to the following:

1. Will response rate be jeopardized by asking additional and somewhat different questions?
2. Which of the several follow-up methods is best to use?
3. Are there geographical differences in response rate to specific mail survey procedures?

Method

The study consisted of a $2 \times 4 \times 8$ factorial, fixed-effects field experiment. The first experimental factor, questionnaire length, consisted of two levels: a short, four-page questionnaire consisting of 23 items related to the socioeconomic–demographic profile of the respondent, and a long, six-page questionnaire consisting of the same 23 items plus 26 additional items measuring respondent attitudes toward the telephone. The average time to complete the questionnaire was about 10 min for the short version and 18 min for the longer version.

The second factor was type of follow-up method. Four follow-up methods were selected based on cost and feasibility considerations. Following the mailing of a postcard alert and the questionnaire, the four follow-up methods utilized were as follows: (a) postcard reminder and a second mailing of the questionnaire; (b) telephone reminder and a second mailing of the questionnaire; (c) letter alert followed by a telephone interview to complete the original mail survey on the phone; and (d) same as (c) but without letter alert.

The third factor, geographical heterogeneity, consisted of eight distinct geographical regions of the country. The geographical areas were chosen based on a cluster analysis of 100 geographical areas representing the total population for the Bell System telephone customers. Each geographical cluster consisted of homogeneous customers with respect to both their socioeconomic–demographic profiles and their telephone behavior profiles. The following eight geographical areas were chosen to represent each of the eight clusters: southeast Massachusetts, San

The authors wish to acknowledge their appreciation to R. B. Ellis and N. J. Mammana of the American Telephone & Telegraph Company and to Lee J. Cobb of Crossley Surveys, Inc. who performed the field work.

Requests for reprints should be sent to Jagdish N. Sheth, 146 Commerce West, University of Illinois, Urbana, Illinois 61801.

TABLE 1

RESPONSE RATE (%) AS A FUNCTION OF FOLLOW-UP METHODS AND GEOGRAPHICAL DIFFERENCES

Market heterogeneity	Follow-up methods				Average	n
	Postcard reminder	Telephone reminder	Alert and telephone interview	Telephone interview		
Alabama	56.7 (66)	68.2 (66)	47.8 (66)	40.3 (66)	53.2	264
Southeast Massachusetts	64.9 (74)	74.3 (74)	56.2 (74)	61.6 (74)	64.2	296
Suburban Chicago	74.6 (66)	75.6 (66)	40.6 (66)	63.2 (66)	63.5	264
Fort Worth	78.8 (66)	76.8 (66)	67.2 (66)	87.5 (66)	77.6	264
Eastern Wisconsin	71.2 (66)	87.9 (66)	54.5 (66)	61.5 (66)	68.8	264
Philadelphia	53.0 (66)	72.3 (66)	60.6 (66)	74.2 (66)	65.1	264
Arizona	76.9 (66)	75.8 (66)	68.2 (66)	78.8 (66)	74.9	264
San Jose	80.9 (66)	80.5 (66)	60.9 (66)	60.3 (66)	70.7	264
Average	69.6	76.4	57.0	65.9	67.3	

Note. Cell ns are given in parentheses.

Jose, Arizona, Philadelphia, eastern Wisconsin, suburban Chicago, Fort Worth, and Alabama.

Within each of the eight markets, a representative random probability sample of 264 telephone customers was drawn and equally divided among the eight experimental conditions created by the two questionnaire lengths and the four follow-up procedures. Thus, each experiment cell had a sample of 33 customers. The only exception was with respect to the eight experimental groups in the southeast Massachusetts region where a sample of 296 customers was drawn and equally divided into the eight experimental conditions to compensate for the anticipated higher rate of disconnection of telephone service due to the end of the summer season.

The mailing dates including those for the follow-up procedures were designed in such a way as to compensate for the regional variations in postal delays. Four attempts were made to contact each respondent for the telephone reminder and the telephone interview procedures. Finally, calls were made at various times of the day and evening which were considered most appropriate based on past experiences of the survey company. In short, every possible effort was made to minimize the differences in the situational factors (time, person, and opportunity to reach) in the field experiment setting between the telephone and the mail interviews. However, it is possible that these efforts may themselves have contributed to some extent in producing differences among experimental cells.

RESULTS

The experimental results are summarized in Table 1. The range of the rate of response varied from a low of 40.3% in Alabama for the telephone interview to a high of 87.9% in eastern Wisconsin for the telephone reminder.

Analysis of variance was performed on the data utilizing Tukey's (1949) additivity model since each cell had only one observation. Furthermore, following Tukey (1949), the test for nonadditivity (Winer, 1971, pp. 394–397) was performed on the data which indicated that the joint effects of all these factors could be treated as an estimate of the experimental error, $F(1, 20) = .11$. Based on the analysis of variance, it was concluded that:

1. The questionnaire length has no significant effect on the rate of return, $F(1, 63) = .04$. The longer questionnaire apparently did not act as a deterrent to the rate of response.

2. The follow-up methods had a significant effect ($p < .001$) on the rate of response, $F(3, 63) = 16.94$. The telephone reminder worked best and the telephone interview with alert had the worst effect.

3. The geographical differences were significantly responsible ($p < .001$) for the differential rate of response, $F(7, 63) = 7.44$. The best response came from the Fort Worth area and the worst response from the Alabama area.

4. There was no interaction effect between the questionnaire length and the follow-up methods, $F(3, 63) = .37$. Thus, the significant effects of follow-up methods are independent of the questionnaire length.

5. The questionnaire length, however, was marginally related ($p < .10$) to geographical differences. Certain geographical areas tended to respond better with shorter questionnaires, and

others tended to respond better with longer questionnaires, $F(7, 63) = 2.20$.

6. There was a significant interaction effect ($p < .05$) between the follow-up methods and geographical differences, $F(21, 63) = 2.20$. Some follow-up procedures work better in some geographical areas, and others work better elsewhere. This clearly suggests the need for a strategy of market segmentation in terms of the utilization of follow-up methods.

In order to localize the source of experimental variance attributed to the two significant factors (follow-up methods and geographical differences), the Neuman-Keuls test was performed on the ranges of pairwise, within-factor differences (Winer, 1971, pp. 191–196). The results clearly indicated that most of the within-facto variance arose due to the two follow-up procedures of the telephone interview with alert and the telephone reminder; the former was significantly worse ($p < .01$) than the other three follow-up methods, and the latter was significantly better than the postcard reminder ($p < .05$) and significantly better than the telephone interview with alert ($p < .01$) as well as without alert ($p < .01$).

In regard to the geographical differences factor, Alabama was significantly poorer ($p < .05$) in generating high response compared to all seven other regions. Similarly, Forth Worth was significantly superior in generating a better response rate than the four markets of Alabama, suburban Chicago, southeast Massachusetts, and Philadelphia ($p < .05$). In fact, nearly three fourths of the variance attributed to the geographical differences factor was localized in the extreme response rates in Alabama and Fort Worth; the other six regions tended to be fairly homogeneous in response rate.

The implications from the above experimental results are somewhat obvious. First, since the length of the questionnaire had no adverse effect, at least within the range tested in the experiment, we should not be overly concerned about its impact on the rate of response (Scott, 1961). However, our range of questionnaire length was small, and the additional questions were attitudinal rather than demographic, making generalization difficult. The change in the nature of the task or questionnaire content may permit an increase in length without lowering response rate. Second, a telephone reminder works best, and, at the same time, is quick, efficient, and easy to control. Since there are no real cost differences between post reminder and telephone reminder for the telephone company or others who have WATS (Wide Area Telephone Service) lines, the added advantage in response rate produced by personalizing the reminder by phone is a clear plus in favor of that type of follow-up procedure.

Third, this study suggested that not all geographical areas of the United States respond the same way. It is therefore necessary to oversample in poorly responding areas in order to compensate for a smaller rate of return. Finally, the most significant implication of this experimental study lies in the interaction effects between the follow-up methods and geographical differences; according to the experimental results, a single procedure does not necessarily generate better results across geographical regions. Thus, it is best to segment the areas and utilize differential follow-up procedures best suited for each geographical segment. From the results summarized in Table 1, it is obvious that while telephone reminders were the best follow-up method in terms of overall impact, the postcard reminder was better in the Arizona area and the telephone interview was better in the Forth Worth area. Accordingly, it may be advantageous to customize follow-up procedures for each type of geographical area to the extent that the cost of differential methods does not become prohibitive. It is surprising to note how little emphasis has been placed on the strategy of segmentation in survey research work.

REFERENCES

Champion, D., & Sear, A. Questionnaire response rate: A methodological analysis. *Social Forces,* 1969, *47,* 335–339.

Erdos, P. L. *Professional mail surveys.* New York: McGraw-Hill, 1970.

Gullahorn, J., & Gullahorn, J. An investigation of the effects of three factors on response to mail questionnaires. *Public Opinion Quarterly,* 1963, *27,* 294–296.

Landy, F. I., & Bates, F. The noneffect of three variables on mail survey response rate. *Journal of Applied Psychology,* 1973, *58,* 147–148.

Linsky, A. A factorial experiment in inducing responses to a mail questionnaire. *Sociology and Social Research,* 1965, *2,* 183–189.

Scott, C. Research on mail surveys. *Journal of the Royal Statistical Society,* 1961, *124,* 143–205.

Tukey, J. W. One degree of freedom for nonadditivity. *Biometrics,* 1949, *5,* 232–242.

Winer, B. J. *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill, 1971.

(Received June 12, 1974)