

Richard A. Kulka, Stewart J. Shirey, R. Paul Moore, and Nelson Woodbury, Research Triangle Institute

For many years in survey methods research investigators of methodological phenomena, like researchers in most substantive areas, have sought to discover and isolate universal principles or "laws," rather than those dependent upon special sets of circumstances and/or unique populations. While such efforts have obviously met with considerable success, in that sophisticated, effective, empirically-based survey methods have evolved from that process, more recently we have noted a healthy and somewhat opposite tendency in methodological research to at least question the general applicability of some of these traditional approaches to survey research in certain subject areas and especially with certain significant or special subgroups of the population. For example, although a controversy still rages on this issue, some have suggested that personal interview surveys may be less effective than telephone surveys in eliciting candid responses to questions on sensitive or socially undesirable topics (e.g., Dillman, 1978; Horton and Duncan, 1978; Sudman and Bradburn, 1974; Rogers, 1976; Locander et al., 1976), while others suggest that mail surveys may elicit more accurate data on certain topics than either of these two survey modes (Semiatiycki, 1979). Similarly, some analyses recently reported by Herzog, Rodgers, and Kulka (1981; Kulka, Herzog and Rodgers, 1981) addressed the appropriateness of interviewing older respondents by telephone.

Within the voluminous mail survey literature as well, the search for universals has been rather intense, but recent literature reviews contain a distinct hint that the relative efficacy of several techniques commonly recommended for enhancing response rates and quality in postal surveys may well depend highly on the particular subject matter addressed and/or the specific population of interest (e.g., Duncan, 1979; Kanuk and Berenson, 1975; Linsky, 1975; Heberlein and Baumgartner, 1978). Thus, in designing a mail survey one must peruse the literature carefully in search of subtle clues as to which approaches are likely to be more or less effective with the specific target population to be surveyed and/or for the designated survey topic.

This is precisely the situation which faced a research team at RTI charged with designing the National Sample Survey of Registered Nurses, the second in a series of surveys designed to collect detailed statistical data on the employment status, distribution, licensure status, and recent educational experiences of registered nurses. Because of the continuing nature of this large-scale survey as a primary method of monitoring nursing resources, and the corresponding need for high quality and cost-efficient data collection methods, the benefits of conducting methodological studies within

the survey to determine which mail survey procedures result in the most optimal combination of maximum response speed, rate, and quality for the least cost within the proposed population are considerable. Specifically, while a number of different factors (some quite trivial) have been shown in the survey methods literature to exert a significant influence on mail survey responses, few such studies have been conducted with professional nurses (or even with similar populations), and the few studies which have been done suggest that factors affecting the responses of other population groups may or may not exert a similar influence on the response behaviors of nurses.

For example, while a number of studies suggest that an advance letter or phone call to respondents before they receive a mail questionnaire will increase response rates (Duncan, 1979; Linsky, 1975), Kephart and Bressler (1958), in a mail survey of professional nurses, found a preliminary letter to be very ineffective, either in isolation or in combination with a followup letter, a finding which might be applicable to many fairly homogeneous populations (Parsons and Medford, 1972). Similarly, while a number of studies suggest that the use of more expensive postage arrangements on outgoing envelopes results in a higher proportion of returns (Kanuk and Berenson, 1975; Linsky, 1975), in the same study Kephart and Bressler (1958) found a distinct advantage for special delivery over first class mail, but no significant advantage for airmail over first class postage.

Consistent with this latter finding, in an experiment conducted within the 1977 National Sample Survey of Registered Nurses (American Nurses' Association and Westat, Inc., 1978), comparing hand-stamped return envelopes with business reply envelopes, final returns were no higher with stamped envelopes (using either conventional or commemorative stamps) than with postage-permit return envelopes, a finding once again in contrast with much of the general literature (Duncan, 1979; Linsky, 1975). However, while the same experiment found that inclusion of a regular sharpened No. 2 "Thank You" pencil as an incentive had essentially no effect on the ultimate response rate, the enclosure of a cash reward as small as a quarter has significantly increased response rates to questionnaires in studies of professional nurses (Kephart and Bressler, 1958), presumably more because of its symbolic than monetary value. Follow-up letters and postcards, personalization of correspondence, and stressing the importance of the respondent to study success--all three consistently found effective in increasing response rates to mail surveys (Dillman, 1978; Duncan, 1979; Linsky, 1975)--have likewise been used effectively in studies of professional nurses (Kephart and Bressler,

1958; Linsky, 1965).

From this brief overview of the limited evidence on mail survey response garnered specifically from studies of professional nurses, the most logical conclusion is that, with few exceptions, we currently know very little about how to assure maximum response among this particular group, and that the uncritical adaptation of several techniques conventionally touted in the general mail survey literature as effective in increasing response rates may well result in a substantial amount of wasted effort and resources in future studies of the nursing profession. To avoid such waste--to achieve maximum response at the lowest possible cost--systematic methodological research has been used as a component of the National Sample Survey of Registered Nurses to critically evaluate the effectiveness of each major component of the survey design in producing both timely and adequate response and high quality data.

Method

Design of the Experiment

In developing a specific experimental design to accomplish this objective, we were guided by two major criteria. First, we sought to test a set of procedures which in combination added no additional costs to the survey as originally proposed, by balancing procedures which involved additional costs with others which resulted in at least equivalent cost savings. Second, we sought to develop a total design which, in our professional judgment, would result in at least as high a response rate as the set of non-experimental procedures originally proposed for the study. Our proportional allocation of the sample to each of the experimental conditions was designed to carefully balance these two criteria while also partitioning the total sample in a way which provided an adequate number of cases in each cell to allow for multivariate analysis of the experimental factors.

First, our decision to vary the use of an advance letter was predicated on the assumption that this procedure, although proposed in the original study design, may not be particularly effective in stimulating the responses of registered nurses. Thus, we did not expect those not receiving the advance letter to respond substantially different from those who did. In combination with other factors, however, it was possible that the use of an advance letter might be effective. Among those receiving a questionnaire mailed at bulk-rate, for example, advance notification may serve to stimulate response to a level comparable to or greater than that achieved among those who receive a questionnaire by first class mail.

A second major factor--the personalization of correspondence--was not proposed in the original design because of its high cost, but since the mail survey literature (e.g., Dillman, 1978; Duncan, 1979; Linsky, 1975), including Linsky's (1965) study of nurses, generally provides support for the effectiveness of that approach in increasing response

rates, we chose to evaluate its benefits relative to these increased costs at two phases of the survey--for the advance letter and the initial questionnaire mailing--in a more modest way with the latter than the former because of differences in cost. Third, as implied in our discussion above, we mailed a substantial proportion of questionnaires by non-profit bulk-rate rather than first class, at a substantial cost savings, primarily because much of the research literature, including a recent study of professional psychologists and psychiatrists (Kahle and Sales, 1978), suggests that these alternative forms of mailing may not result in different rates of response.

Fourth, although evidence from the initial National Sample Survey of Registered Nurses suggested that the enclosure of a pencil as an incentive would not prove effective in stimulating responses to the survey, we included a No. 2 pencil for half the sample in this second survey as an aid. Because the survey questionnaires were designed for processing by optical mark scanning equipment, the provision of a pencil in this context had the potential for improving the quality if not the quantity of response in this survey. Finally, although prior research at RTI had established the effectiveness of mailgrams in the latter stages of mail survey follow-up, research by Dillman (1978, Dillman et al., 1974) and others (e.g., House et al., 1977; Heberlein and Baumgartner, 1978) suggests that special delivery and certified mail (both less expensive) may prove equally effective (as may first-class postage).

Implementation of the Design

The survey instrument was originally mailed to a sample of over 40,000 registered nurses nationwide which were randomly pre-assigned to the survey conditions described below, with the proportions assigned to each condition contingent on the cost and projected outcome considerations described above. After duplicates, undeliverables, and other ineligible respondents (e.g., the deceased) were eliminated from the sample, a total of 37,835 nurses were subject to these experimental conditions.

Advance letters. Approximately equal numbers of respondents (one-third) received either no advance letter, a personalized advance letter, or a non-personalized advance letter. Those receiving a non-personalized advance prenotification were sent a form letter with the salutation "Dear Registered Nurse" in a No. 10 envelope affixed with a mailing label. In contrast, nurses receiving a personalized advance mailing received a letter containing the same salutation, but the mailing also contained a typed name and address both on the letter and on the outside envelope, rather than a mailing label. However, because these individualized addresses were also computer-generated, the "personalized" advance letters also contained a 6-digit identification number both on the envelope and as part of the inside address, a factor which probably detracted somewhat from

the "personal touch" this method was intended to create.

Type of postage. About two weeks after the advance letters were mailed, approximately 60 percent of the sample were sent questionnaires in envelopes with first-class metered postage and "FIRST CLASS" printed on the envelope. The remaining 40 percent received a third-class mailing, in envelopes with "bulk-rate" printed in the upper right-hand corner. Both types of mailings used window envelopes, designed so that the label affixed to the questionnaire would show the appropriate name and address through the window, and the envelopes also contained a business reply envelope and a cover letter. A small subset of the first-class mailings were also "personalized"--the cover letter included a typed address, a white envelope with no window was used, and the nurse's name and address were typed on the envelope.

Inclusion of pencil. A No. 2 pencil was included in half of the envelopes mailed across all treatments, with no mention of the pencil included in the cover letter.

Interim procedures. All sample members received a "thank you/reminder" postcard about three weeks after receiving the questionnaire. The postcard went out automatically regardless of whether or not a nurse had responded or what treatment combination he or she was in. About two weeks after the postcard, a second questionnaire was mailed first-class to all nonrespondents in a mailing which included no experimental manipulations.

Third follow-up. For the third follow-up, mailed approximately six weeks after the second follow-up package, nonrespondents were randomly allocated to four experimental groups independent of earlier treatment assignments. About one-fifth received a questionnaire sent via special delivery and another 20 percent by first-class mail. Half of the remaining nonrespondents were sent questionnaires by certified mail, with the other 30 percent receiving mailgrams without questionnaires. Except for the mailgrams, each envelope, regardless of postage type, included a cover letter, questionnaire, and postage-paid return envelope.

Results

Response Rate

This series of mailouts, commencing with the initial mailing and concluding with the special third follow-up mailing, resulted in 30,433 questionnaires being returned for an overall response rate of just over 80 percent. In calculating this rate, those who had no opportunity to respond--undeliverables and the deceased--were deleted from the original sample, along with duplicates, leaving a total of 37,835 eligible respondents.

Response rates by experimental condition are presented in column (1) of Table 1, which, in spite of the unusually large number of cases in most of the groups being compared, shows (with one notable exception) surprisingly little variation in the final

Table 1
Relationship of Experimental Factors to Rate and Completeness of Survey Response

Experimental Factor	Response Factor			
	(1) Response Rate		(2) Response Completeness	
	%	N ^b	%	N ^c
A. Prenotification				
Advance letter				
Personalized	81.1	12,041	6.2	9,760
Not personalized	81.3	12,015	6.0	9,764
No advance letter	79.2	13,779	5.9	10,909
B. Postage/Personalization				
First class postage				
Personalized	78.1	1,116	6.3	872
Not personalized	80.4	21,790	6.0	17,510
Third-class (bulk-rate) postage	80.7	14,929	6.1	12,051
C. Incentive/Aid				
Pencil	80.6	18,959	6.3	15,288
No pencil	80.2	18,876	5.8	15,145
D. Third Follow-up postage				
First class mail	22.9	1,939	7.2	443
Special delivery	31.0	1,831	8.3	568
Certified mail	41.0	2,627	7.4	1,078
Mailgram	18.0	2,800	8.1	504

^a Percentages with one or more of three selected items blank or double-marked.

^b Based on number of questionnaires which could have been returned (i.e., duplicates and undeliverables are omitted).

^c Based on number of questionnaires returned.

response rates induced by these procedures, despite the fact that each of these manipulations has been effective in increasing response rates in at least a few, and in most cases several, prior mail survey experiments. An analysis of variance which summarizes the main and interactive effects of the first three conditions on response rates is presented in the first column of Table 2.

Although based on our review of prior research on prenotification (including one study involving the same basic population) we had hypothesized that the use of an advance letter would not be effective in increasing the response rate in this study, nurses receiving an advance letter did respond in somewhat greater numbers than those who did not receive advance notification. However, while the difference is clearly significant in a statistical sense, the practical significance of a difference of only 2 percent in light of the substantial additional costs (in excess of \$7,000) associated with achieving that difference in a survey of this size is questionable. Moreover, whether or not an advance letter is personalized, at least in the limited manner implemented in this survey, apparently makes little or no difference in the final response rate.

Second, consistent with a recent study of other professionals reported by Kahle and Sales (1978), and in contrast with studies conducted on more heterogeneous populations (e.g., Kernan, 1971), using first-class

Table 2
Analysis of Variance Summary
for Response Rate and Response Speed

Source of Variation	Response Factor					
	(1)			(2)		
	Response Rate			Response Completeness		
	df	MS	F	df	MS	F
Main Effects						
A. Prenotification	2	1.747	11.11**	2	0.027	0.48
B. Postage/Personalization	2	0.364	2.31	2	0.006	0.11
C. Incentive/Aid	1	0.155	0.99	1	0.172	3.05
Interaction Effects						
A X B	4	0.142	0.90	4	0.091	1.61
A X C	2	0.042	0.27	2	0.012	0.22
B X C	2	0.047	0.30	2	0.042	0.74
A X B X C	4	0.255	1.62	4	0.028	0.24
Error	37,817	0.157	--	30,415	0.565	--

* p < .05

** p < .01

postage on the outgoing questionnaire mailing was not more effective than bulk-rate envelopes in stimulating response rates, and the practical significance of this non-significant difference is far greater than the small effect noted for prenotification, since the cost differential between the non-profit bulk rate and first-class postage is substantial. Thus, in spite of a preponderance of conventional wisdom to the contrary, "higher class" outgoing postage does not necessarily result in higher response rates, perhaps especially among a population of professionals. Moreover, a "personalized" first-class mailing was no more effective than the standard non-personalized questionnaire mailing in enhancing response rates.

Third, although pencils have apparently represented effective incentives for enhancing responses in other populations (e.g., Pucel et al., 1971), consistent with results obtained in the first National Sample Survey of Registered Nurses (American Nurses' Association and Westat, Inc., 1978) inclusion of a pencil had no effect whatsoever on the response rate.

Although, as noted earlier, it was conceivable that one or more of these factors might have been more or less effective in combination with the others, Table 2 indicates no significant interactions among these conditions.

In contrast, results obtained from the third follow-up mailing indicate some dramatic differences in response among groups receiving different types of postage. Differences among all of these treatments are statistically significant, with certified mail clearly producing the best follow-up

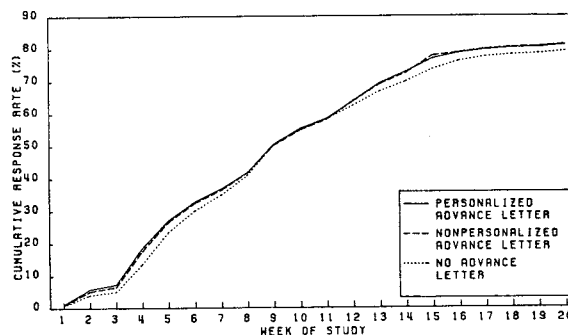
results in this particular population. In particular, the use of certified mail was more than twice as effective as a mailgram, and even an additional first-class mailing proved to be more effective than a mailgram among these professional nurses, although the latter was the most expensive treatment of the four implemented. Although certified mail was the second most expensive procedure used (costing slightly more than special delivery in this application), its clear superiority in inducing reluctant respondents to return a questionnaire still rendered it the most cost-effective. In effect, certified mail was as effective in stimulating response among this group of persistent nonresponders as our initial mailout was for the total sample (about 40 percent in both cases).

The poor performance of the mailgram in this study was somewhat surprising, in that previous research experience at RTI with the National Longitudinal Study of the Class of 1972 had led us to expect the mailgram to produce a beneficial effect on overall response, equal to or greater than that produced by certified mail or special delivery. One possible explanation for the rather poor results obtained with the mailgram is that this was the only condition which did not make another questionnaire available to the nurse, although evidence on the need to include a questionnaire in a follow-up mailing is mixed (Futrell and Lamb, 1981; Heberlein and Baumgartner, 1981). Evidently, however, the certified mail and special delivery treatments had the intended positive effect of emphasizing the importance of the study to these nonresponding nurses, an effect apparently strong enough to overcome some possible negative factors, such as the necessity, in some cases, to go to the post office to sign for the certified mailing.

Response Speed

In addition to final responses rates, we sought to investigate interim response rates or response speed--the time lag between the mailing of the questionnaire to our respondents and its return to RTI--a factor which along with response quality has only recently been of much concern in the mail survey literature (e.g., Houston and Ford, 1976).

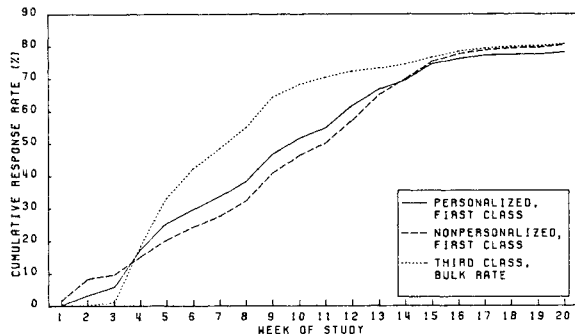
FIGURE 1. PRENOTIFICATION RELATED TO SPEED OF RESPONSE



Figures 1 to 3 present cumulative response distributions by week for the first three

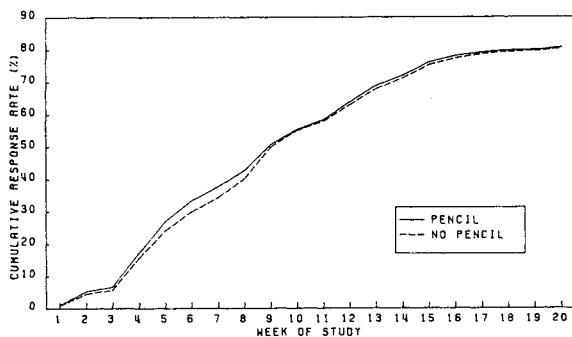
conditions previously examined in Table 1. The results in each case are fairly consistent with what one might expect. The slight difference in response between those receiving and not receiving an advance mailing are consistent throughout the survey period (Figure 1), and although those receiving personalized and non-personalized advance letters tend to criss-cross repeatedly throughout the survey period, their interim response rates never differ by much.

FIGURE 2. POSTAGE/PERSONALIZATION RELATED TO SPEED OF RESPONSE



In contrast, since bulk-rate mailings by definition take longer to reach respondents than questionnaires mailed with first-class postage, the latter have a clear speed of response advantage for the first few weeks of the survey, an advantage which is clearly overcome, however, and exceeded by the bulk-rate group within four weeks. Having gained an advantage, the bulk-rate group does not relinquish it until very late in the survey, essentially only after the third follow-up mailing. Among the first class mailings, personalized questionnaire mailings, while having no advantage in producing better final response rates, do apparently elicit somewhat quicker responses after the first couple of weeks of the survey.

FIGURE 3. INCENTIVE/AID RELATED TO SPEED OF RESPONSE



Finally, although the pencil "incentive" has no effect on final response rate, respondents receiving a pencil appear to respond slightly faster than those who must use their own.

Response Completeness

To investigate potential effects of the experimental conditions on response completeness, two measures were created. One, speci-

fic to the pencil/no pencil condition, tallied the number of questionnaires completed in ink. Since the questionnaires were to be read by optical mark sensing equipment, a questionnaire completed in ink had to be completely remarked by pencil in-house prior to scanning. Overall, only 913 (2.4 percent) questionnaires were completed in ink, a smaller number than we had anticipated, 74 percent of which had been mailed without a pencil. Specifically, while only 1.3 percent of the questionnaires mailed with a pencil were filled out in ink, 3.6 percent of those sent without a pencil were. Thus, those not receiving a pencil were almost three times as likely as those mailed a pencil to respond in ink, but the overall number using ink was so small in this survey that inclusion of the pencil proved not to be cost-effective.

A second and equally crude "completeness" criteria was derived by counting the number of blank or multiple-marked items on three questions from different sections of the questionnaire. The use of a larger number of questions was not possible because extensive editing was done on most items prior to optical scanning. Overall, only six percent of the responding nurses gave incomplete data on one or more of these three items. As indicated in column (2) of Table 1 and the analysis of variance summary in Table 2, there are no significant differences among experimental conditions in response completeness as indicated by this highly-skewed variable. In the future, however, we intend to conduct some additional analyses on response completeness and quality with these data. One hint of a possible direction is evident in Table 1. Note that respondents to the third follow-up are somewhat more likely to give incomplete responses than those responding earlier, a finding consistent with earlier research suggesting that late or reluctant responders may provide data of lower quality than early or willing respondents (Cannell and Fowler, 1965).

Conclusion

Although the data presented in this paper are obviously somewhat limited, in that they derive from one study of a fairly special population and reflect only a handful of potential response-inducing factors, the results described suggest at least some caution in assuming that even procedures which have received broad-based research support in the mail survey literature are readily generalizable to populations of interest other than those surveyed in the methodological studies which served to establish the efficacy of those techniques. While on the surface this general caution would appear self-evident and, in theory, would likely receive widespread endorsement among survey practitioners, in our view it has not as yet been well assimilated into the mail survey research enterprise.

Moreover, in reflecting upon the results presented here it is important that we not ourselves undermine or ignore this general principle by attempting to generalize the

differences reported here (or lack thereof) to the full range of populations that we might be called upon to survey in the future. Rather, the most valuable general lesson or orientation to be gained from the evidence presented is, in our view, a healthy skepticism and caution in selecting a set of design features intended to ensure maximum response in a mail survey of a particular population. By proceeding with such caution, by carefully weighing and evaluating techniques commonly touted in the research literature against the specific characteristics of the population to be surveyed, and subjecting such techniques to direct empirical assessment wherever practicable, it should be possible to achieve further improvements in the overall effectiveness of mail surveys in a wide variety of applications, while also potentially saving a great deal of unnecessary difficulty and expense, perhaps in many more applications than we might currently believe.

References

- American Nurses' Association, and Westat, Inc. 1977 National Sample Survey of Registered Nurses - (Vol. III): Technical Report, Appendices to Field Procedures. Kansas City, MO: American Nurses' Association, 1978.
- Cannell, C. F., and Fowler, F. Comparison of hospitalization reporting in three survey procedures. Vital and Health Statistics. Washington, DC: U.S. Public Health Service, Series 2, No. 8, 1965.
- Dillman, D. A. Mail and Telephone Surveys: The Total Design Method. New York: John Wiley, 1978.
- Dillman, D. A., Christenson, J. A., Carpenter, E. H., and Brooks, R. M. Increasing mail questionnaire response: A four-state comparison. American Sociological Review, 1974, 39, 744-756.
- Duncan, W. J. Mail questionnaires in survey research: A review of response inducement techniques. Journal of Management, 1979, 5(1), 39-55.
- Putrell, C. M., and Lamb, C. W., Jr. Effect on mail survey return rates of including questionnaires with follow-up letters. Perceptual and Motor Skills, 1971, 52, 11-15.
- Heberlein, T. A., and Baumgartner, R. Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature. American Sociological Review, 1978, 43, 447-462.
- Heberlein, T. A., and Baumgartner, R. Is a questionnaire necessary in a second mailing? Public Opinion Quarterly, 1981, 45(1), 102-108.
- Herzog, A. R., Rodgers, W. L., and Kulka, R. A. Interviewing older adults: A comparison of telephone and face-to-face modalities. Paper submitted for publication, 1981. (Revised version of three papers presented at the 33rd Annual Meeting of the Gerontological Society, San Diego, November 1980).
- Horton, R. L., and Duncan, D. J. A new look at telephone interviewing methodology. Pacific Sociological Review, 1978, 21(3), 259-274.
- House, J. S., Geber, W., and McMichael, A. J. Increasing mail questionnaire response: A controlled replication and extension. Public Opinion Quarterly, 1977, 41, 95-99.
- Houston, M. J., and Ford, N. M. Broadening the scope of methodological research on mail surveys. Journal of Marketing Research, 1976, 13, 397-403.
- Kahle, L. R., and Sales, B. D. Personalization of the outside envelope in mail surveys. Public Opinion Quarterly, 1978, 42, 547-550.
- Kanuk, L., and Berenson, C. Mail surveys and response rates: A literature review. Journal of Marketing Research, 1975, 12, 440-453.
- Kephart, W. M., and Bressler, M. Increasing the responses to mail questionnaires. Public Opinion Quarterly, 1958, 22, 123-132.
- Kernan, J. B. Are bulk-rate occupants really unresponsive? Public Opinion Quarterly, 1971, 35, 420-422.
- Kulka, R. A., Herzog, A. R., and Rodgers, W. L. On the appropriateness of interviewing older respondents by telephone. Paper presented at the annual meeting of the American Association for Public Opinion Research, Buck Hill Falls, Pennsylvania, May 1981.
- Linsky, A. S. A factorial experiment in inducing responses to a mail questionnaire. Sociology and Social Research, 1965, 49, 183-189.
- Linsky, A. S. Stimulating responses to mailed questionnaires: A review. Public Opinion Quarterly, 1975, 39, 82-101.
- Locander, W., Sudman, S., and Bradburn, N. An investigation of methods, threat, and response distortion. Journal of the American Statistical Association, 1976, 71, 269-275.
- Parsons, R. J., and Medford, T. S. The effect of advance notice in mail surveys on homogeneous groups. Public Opinion Quarterly, 1972, 36, 258-259.
- Pucel, D. J., Nelson, H. F., and Wheeler, D. N. Questionnaire follow-up returns as a function of incentives and respondent characteristics. Vocational Guidance Quarterly, 1971, 19, 188-193.
- Rogers, T. F. Interviews by telephone and in person: Quality of responses and field performance. Public Opinion Quarterly, 1976, 40, 51-65.
- Siemiatycki, J. A comparison of mail, telephone, and home interview strategies for household health surveys. American Journal of Public Health, 1979, 69(3), 238-245.
- Sudman, S., and Bradburn, N. M. Response Effects in Surveys: A Review and Synthesis. Chicago: Aldine, 1974.