Factors Influencing Physicians' Response to Mailed Questionnaires

WARREN H. GULLEN, MD, MPH, and GLEN E. GARRISON, MD

IN MAILED questionnaire surveys, epidemiologists have usually used questionnaires designed and reproduced to be as attractive as possible, accompanied by the most finished or professional-appearing covering letters and sent by first-class mail. These practices are based on the assumption that appearance and type of postage will lead to higher response rates than would be achieved otherwise. This assumption results partly from intuitive reasoning and partly from the extensive literature on the use of mail surveys.

These previously reported investigations, many of which were made by commercial firms (for example, product-related surveys), have yielded incomplete data that are often difficult to interpret. For instance, one study of random samples of a general population (1) showed that adhesive postage stamps produced

a slightly better response than bulk postage. A consumer survey of a random sample of households revealed no significant difference between the response to printed and mimeographed questionnaires (2). A study of automobile owners (3) did not elicit a significant difference in responses between metered postage

Dr. Gullen is associate professor and chief, epidemiology section, department of community medicine, and Dr. Garrison is professor and chairman, department of community medicine, and medical director, division of continuing education, Medical College of Georgia. Tearsheet requests to Warren H. Gullen, MD, Medical College of Georgia, Augusta, Ga. 30902.

and the ordinary adhesive type. Surveys of selected college graduates (4) and nurses (5) revealed that higher classes of postage may produce a higher yield, although the difference may be rather small. Another study of nurses (6) showed that a letter with a personal salutation and signature got a better response than one with a mimeographed salutation and signature. In an investigation of a random sample of women (7), the only major difference in response observed was related to the age of the woman (the response decreased with increased age). A mailing to teachers (8) produced no significant difference in response by modes used.

No attempt has been made in the preceding discussion to evaluate analytically the studies cited because they were of different degrees of sophistication and were designed in different ways to assess the effects of various factors on the response of various target populations. Furthermore, none were directly relevant to the question that was the basis for the study reported in this article because they were not assessments of responses of practicing physicians to different types of mailings from a medical school; nor did they use the same treatment factors in the same combinations as were included in our study. We have tested the basic assumption and its relevance to practicing physicians and have investigated other factors influencing the rates of response to mailed questionnaires.

Methods

The section of epidemiology and the division of continuing education of the Medical College of Georgia conducted a survey to assess the attitudes of physicians in private practice regarding certain common administrative practices in continuing education. Concomitantly, an experiment was executed to determine which of several kinds of questionnaire mailings would stimulate the highest response rate. The questions in the questionnaire were substantive, simple, relevant to almost all physicians, and not concerned with sensitive or controversial subjects. For example, two questions concerned the preference of the physician for various types of formats for announcements of meetings or courses and the time interval that should be allowed between the mailing of course brochures and the beginning of the course itself.

The two main treatment variables in this experiment were format (including both design and reproduction process) and postage (primarily first class compared with bulk or third class). Specifically, there were seven treatments as follows:

1. Mimeographed two-part tear-off post card sent by bulk mail. There was no business heading, personal salutation, or date. The medical director of continuing education had obviously signed the mimeograph stencil.

2. Printed two-part post card sent by bulk mail. There was no business heading, personal salutation, or date. Director had signed the offset master.

3. Same as treatment 2 except postage was first class and metered (instead of the printed permit used in treatment 2).

4. Mimeographed form letter sent by bulk mail. There was no heading, personal salutation, or date. The director had signed the stencil. A return envelope in business-reply style was enclosed.

5. Mimeographed form letter sent by first-class mail. (Same as 4 except that the postage was first class and metered instead of the printed permit used in treatment 4.)

6. Business-style letter sent by metered first-class mail. The letters were reproduced on letterhead bond by the multilith process. The letter had an individually typed business heading, personal salutation, and a date. The entire letter, including the director's signature, was produced in a way that would require close inspection to determine that it had not been individually prepared for the recipient. A business-reply return envelope was enclosed.

7. Identical to treatment 6 except the outgoing postage was an adhesive stamp instead of metered postage.

Return postage was prepaid in all treatments, and the text of the letter and questionnaire was the same for all seven.

The treatments were assigned in sequence to Georgia physicians in private practice by using a 1967 American Medical Directory. That is, the first physician with type of practice code 01 (full-time general practice or other full-time specialty practice) or 02 (general practice with some specialty practice) listed in the Georgia section of the directory was assigned treatment 1; the next such physician, treatment 2; the next, treatment 3; and so on through treatment 7. The sequence was then repeated starting with treatment 1, and the procedure was continued until all such physicians listed in the Georgia section had been allocated a treatment category. Each physician was given an identification number, which was also

imprinted on the corresponding questionnaire.

All the questionnaires were mailed on April 10, 1970, and the cutoff for return was May 25, 1970. A "do not forward—return to sender" statement appeared on the front of the envelope or post card mailed to each physician.

The data in the analyses were derived from the returned questionnaires and also from the AMA directory that was used in the treatment allocations. These data were transferred to punchcards and analyzed with the help of a computer.

Results

There were 3,002 questionnaires mailed. Of those, 2,648 (88.2 percent) apparently reached the physicians; the rest (354) were returned by the post office or by others for various reasons (for example, physician had died or had moved). The number of questionnaires counted as returned by the physicians in the analyses was 1,194. Virtually all the returned questionnaires were substantially completed except for a number that were not signed by the physician. Fifteen others were returned but not included in the analyses (12 of them were received after May 25, the cutoff date).

Of the returned questionnaires that contained the date of completion, about 88 percent were completed within 1 week of the mailing, an observation not surprising to those experienced in mail surveys. There was no apparent pattern of differences in speed of completion by the physicians receiving the various treatments.

The percentage of returned questionnaires signed by the physician (signature was optional) was about the same for every treatment except treatment 1. The percentage signed was considerably lower for treatment 1 (about 75 percent compared with 83–88 percent for the rest). This difference was of borderline significance ($X^2_6 = 13.30$, $P < 0.043$).

The return rates for the specific treatments are presented in table 1. A chi-square omnibus test of the hypothesis that all treatment groups in the universe would have equal response rates yielded a X^2_6 of 71.89 and a probability value $<<0.0005$. If the treatments are ordered according to increasing response rate, and tests of significance of differences in proportion (using normal curve approximation) are performed on all two-treatment contrasts composed of adjacent

treatments, it is observed that the only two such tests that are significant are those for treatment 2 compared with treatment 3 ($Z = 2.43$, P (one tail) $= 0.0075$) and treatment 5 compared with treatment 7 ($Z = 2.00$, P (one tail) $= 0.0233$). Z equals standard error units of the unit normal curve. All the rest of the Z values were less than unity. It should be noted, of course, that we tested multiple hypotheses and a posteriori hypotheses. The treatments were easily grouped into three categories according to response rate. The low response group included treatments 1 and 2; the intermediate response group included treatments 3, 4, and 5; the high response group included treatments 6 and 7. This conclusion was strengthened by the observation that responses to treatment 3 were not significantly different from those to treatment 5 ($Z = 1.19$, P (one tail) $= 0.1170$), and that all samples were substantial and of almost equal size.

It is obvious that a number of factors are confounded in the various treatment groups. It is possible, however, for the reader to derive some idea of the effect of some single factors by comparing treatment results, bearing in mind the statements previously given regarding the significance of differences. For example, if one wanted to assess the effect of adhesive stamps compared with metered postage, he would observe the results of treatments 6 and 7, which were not "significantly" different in response rate and are identical to each other except that in treatment 7 an adhesive stamp was used instead of metered postage.

Many comparisons can be made, and the reader can select

Table 1. Responses of physicians in Georgia to mail survey

Treatment number	Number sent	Number returned	Percent returned
1	371	127	34.23
2	377	130	34.48
3	386	166	43.01
4	367	164	44.69
5	385	182	47.27
6	382	218	57.07
7	380	207	54.47
Total	[1] 2,648	1,194	45.09

[1] 3,002 questionnaires were mailed, but 354 were returned by the post office or others for various reasons (for example, physician had died or had moved).

those that are pertinent for him, using the figures in the tables. Unfortunately, a complete factorial design that could be used to investigate all factors and factor combinations would have required a much larger number of treatments, thereby seriously reducing the number of persons in each treatment group. Also, we deemed it unrealistic to include certain combinations. For example, none of the treatments included the offset letter with personal salutation and signature but with bulk postage, because we believed that if an investigator intended to "go first class" he would "go first class all the way." Thus, we decided that the treatment combinations used were the most realistic ones for actual or potential usage. Therefore, it may well be that it is the combinations that are important, and it is not essential to sort out the effects of the individual treatment components.

Possibly, the conclusions regarding the effectiveness of the

Table 2. Analyses of data stratified according to indicated variables

Stratification variable	Significant difference between strata?	Significantly similar patterns within the various strata?	Remarks
Community size	No, $X^2_{28}=36.32$; $0.10<P<0.20$	Yes, $W=0.74$, significant at 1 percent	Ranks of R_j identical to ranks of rates in table 1.
Geographic areas	No, $X^2_{42}=50.26$; $P>0.10$	Yes, $W=0.48$, significant at 1 percent	Ranks of R_j for treatments 6 and 7 were 1 and 2.
AMA membership	Yes, binomial probability $=0.008$	?, $W=0.77$, Spearman $r_s=0.54$, not significant at 5 percent	Few physicians were not AMA members. AMA members had higher response rate. Treatments 6 and 7 significantly better than others among nonmembers also $(Z=1.82)$.
Year of birth	Yes, $X^2_{35}=75.08$; $P<0.005$	Yes, $W=0.64$, significant at 1 percent	Response rate tends to decrease with age. Ranks of R_j almost identical to ranks of rates in table 1. Treatments 6 and 7 best.
School of medical degree	Yes, $X^2_{35}=54.27$; $0.01<P<0.02$	Yes, $W=0.82$, significant at 1 percent	Medical College of Georgia graduates had highest rate. Ranks of R_j identical to ranks of rates in table 1.
Year of medical degree	Yes, $X^2_{42}=83.11$; $P<0.025$	Yes, $W=0.66$, significant at 1 percent	Response rate tends to decrease with age of medical degree. Ranks of R_j almost identical to ranks of rates in table 1. Treatments 6 and 7 best.
Specialty boards	Yes, $X^2_{14}=31.11$; $P<0.008$	Yes, $W=0.89$, significant at 1 percent	Nonsurgical and related-board group had highest rate followed by surgical and related, and non-boarded. Ranks of R_j almost identical to ranks of rates in table 1. Treatments 6 and 7 best.
Primary specialty	?, $X^2_{35}=47.89$; $P=$ approx. 0.07	Yes, $W=0.61$, significant at 1 percent	Physicians ranked by specialty in order of decreasing response rates: pediatrics, internal medicine, obstetrics, surgery, general practice. Ranks of R_j essentially identical to ranks of rates in table 1.
Membership in specialty societies	Yes, $X^2_7=33.49$; $P<<0.001$	Yes, $W=0.93$, $r_s=0.86$, almost significant at 1 percent	Members had higher rate than non-members. Treatments 6 and 7 best.
Visiting appointment at medical school	No, $X^2_{14}=15.87$	Yes, $W=0.70$, significant at 5 percent	Ranks of R_j: treatment 7 best, followed by treatment 5 and then treatment 6.

NOTES: W=Kendall coefficient of concordance.
R_j=sum of ranks as used in Kendall's coefficient of concordance.

different treatments may not be valid for every subgroup among the population of private practitioners in Georgia. For that reason, the 2,648 physicians who apparently received the questionnaire were poststratified in various ways—by community size, geographic area, AMA membership, year of birth, medical school, year of medical degree, specialty board, primary specialty, membership in specialty societies, and visiting appointment at a medical school. For each such stratification, the response rate was calculated for each of the multiple treatment-stratum combinations.

Two pertinent questions could be asked about each set of stratified data. First, was there a significant difference in response rates between strata? And, second, did the various treatments tend to evoke similar response patterns within the several strata? To assist in answering the first question, a X^2 was computed within each treatment to compare the rates for the various strata. The seven resultant X^2's were then added, as were their degrees of freedom, to yield an overall X^2.

To provide an answer to the second question, a Kendall coefficient of concordance was computed.

Table 2 contains a tabulation of the major results of these procedures as applied to the various stratifications. It should be pointed out that several of the stratification variables are undoubtedly confounded. It should also be pointed out that in table 2 a question mark (?) appears in a few places instead of a direct statement regarding significance. This question mark indicates that the significance is, in our opinion, equivocal and that the reader should be especially careful to derive his own conclusions (see the probability values).

In summary, the results of this experiment support the assumption usually made by epidemiologists in designing mail surveys that the more attractive and finished the format, the more personal the address, and the higher the class of mail postage, the better the response rate. At least it seems to hold for physicians in private practice, both in the aggregate and in the various substrata.

REFERENCES

(1) Champion, D. J., and Sear, A. M.: Questionnaire response rate: a methodological analysis. Social Forces 47: 335–339 (1969).
(2) Ford, N. M.: Questionnaire appearance and response rates in mail surveys. Journal Advertising Research 8: 43–45 (1968).
(3) Bridge, R. G.: Alternative postage methods in mail surveys. Survey Research Center occasional paper No. 7101. University of California at Los Angeles, January 1971.
(4) Gullahorn, J. E., and Gullahorn, J. J.: An investigation of the effects of three factors on response to mail questionnaires. Public Opinion Q 27: 294–296 (1963).
(5) Kephart, W. M., and Bressler, M.: Increasing the responses to mail questionnaires: a research study. Public Opinion Q 22: 123–132 (1958).
(6) Linsky, A.: A factorial experiment in inducing responses to a mail questionnaire. Sociology and Social Research 49: 183–189 (1965).
(7) Kaplan, S., and Cole, P.: Factors affecting response to postal questionnaires. Br J Prev Soc Med 24: 245–247 (1970).
(8) Mason, W. S., Dressel, R. J., and Bain, R. K.: An experimental study of factors affecting response to a mail survey of beginning teachers. Public Opinion Q 25: 296–299 (1961).

GULLEN, WARREN H. (Medical College of Georgia), and GARRISON, GLEN E.: *Factors influencing physicians' response to mailed questionnaires. Health Services Report, Vol. 88, June–July 1973, pp. 510–514.*

This experiment was executed to determine which of several kinds of questionnaire mailings would elicit the best rate of response from physicians. The two main treatment variables were format (including both design and reproduction process) and postage (primarily first class compared with bulk or third class). The 3,002 physicians in private practice listed in the Georgia section of the 1967 American Medical Association's directory were divided into seven treatment groups. The content of the seven treatments ranged from a crude mimeographed two-part post card sent by bulk mail to an attractive business letter format sent by first-class mail.

The response rates varied from 57 percent for the attractive business letter format sent by first-class mail to 34 percent for the crude mimeographed post card sent by bulk mail. Thus, the results of the experiment support the assumption usually made by epidemiologists in designing mail surveys that the more attractive and finished the format, the more personal the address, and the higher the class of postage, the better the response rate. At least it seems to hold for physicians in private practice, both in the aggregate and in various subgroups.