

FACTORS AFFECTING RESPONSE TO POSTAL QUESTIONNAIRES

SAMUEL KAPLAN, M.D.

AND

PHILIP COLE, M.D.

Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts 02115

Epidemiological studies of chronic diseases often use data gathered individually from many persons, rather than remaining restricted to pre-recorded data. Generally this is done by personal interview, an expensive, time-consuming procedure. The increasing educational level and social consciousness of the population suggest that a carefully designed, highly objective postal questionnaire could substitute for an interview. However, the validity of studies done this way is limited by the response rate, which tends to be lower than in interview studies. Little is known of factors affecting response to postal questionnaires and two studies were undertaken to assess some of these.

METHODS

The first study assessed whether response rate varied with personalization, length of questionnaire or ethnic connotation of the name signed to the letter. The study population consisted of 480 women selected from the 1967 Residents Lists of the 87 towns comprising the contiguous Boston and Brockton, Massachusetts, Standard Metropolitan Statistical Areas. For each decade of age from the 20s to the 70s inclusive, 80 women were chosen at random. Forty women in each decade received a questionnaire on reproductive history, past medical history and demographic data which took 5 to 10 minutes to complete. The other 40 in each decade received a questionnaire on reproductive history which took only 1 or 2 minutes to complete. These are subsequently referred to as the 'long' and 'short' forms respectively. Half of the women in each age decade who received each form were sent a 'personal' letter which included a complete inside address and the salutation 'Dear Mrs. (name)'. The other half of the women were sent an 'impersonal' letter with no inside address and the salutation 'Dear Madam'. In addition, the envelopes sent to women receiving personal letters

and the return envelopes enclosed were stamped, while the corresponding envelopes for the impersonal letters were franked. The body of the personal and impersonal letters was identical. Finally, one-third of the women receiving each questionnaire and each type of letter had their letter hand-signed by a name considered Jewish, a name not associated with any particular ethnic group, and a name that might be recognized as Irish, respectively. As there were 20 women in each length-personalization group within each decade of age, seven women were allocated to each of two signers and six to the remaining signer. The small group was allocated to signers in rotation following an initial choice at random. The title 'M.D.' was also hand-signed in each case, and the signer was identified as 'Project Director'.

A follow-up letter was sent to each woman who had not responded within one month. Each woman's follow-up request was identical in format to the original, only the body of the letter being changed. After another month persistent non-responders were divided into two groups at random. One group received a third letter in the original format via certified mail. The other group was sent letters by regular mail following an attempt by telephone to ask them to cooperate.

The second study assessed whether response rate varied with city of residence, ethnic connotation of the name of the letter signer, use of the title 'Professor' or inclusion of a question asking, in an inappropriate context, the age at which schooling had been completed. The study population consisted of 400 women selected from the Residents Lists of each of the cities of Newton and Worcester, Massachusetts. In each city 100 women were chosen at random for each decade of age from the 20s through the 50s inclusive. Half the women in each city in each decade of age received a questionnaire on reproductive history and the use of oral contraceptives. The other half received the same

questionnaire with an additional question on age when schooling finished. Each of these age, town and questionnaire-specific groups was divided into thirds. One-third received a letter hand-signed by a name not associated with any specific ethnic group without a title, the second third received a letter hand-signed with a name that might be recognized as Irish without a title, and the final third received a letter hand-signed with the same Irish name followed by the title 'Professor'. (Again, the signer to be one letter 'short' in each subgroup was chosen in rotation after an initial random selection.) The letters were otherwise identical. After one month a follow-up letter was sent to non-responders; this was identical to the first letter in format.

RESULTS

Table I shows final response rates for the first study, based on the 436 women deemed eligible for the study. Forty-four women (9.2%) were excluded from the tabulation because their letters were returned undelivered (34) or because an informant advised that the subject was deceased (8) or senile (2). If these women are included the overall response rate would drop from 75.6% to 69.0%. The high exclusion rate was anticipated because, for practical reasons, Residents Lists more than two years old were used. With the exception of 'age of women' the differences shown in Table I are small, and none achieves statistical significance. A trend of decreasing response with increasing age is apparent. Controlling for 'signer' in the analysis, the association of response rate with decade of age is highly significant (summary $\chi^2 = 9.18$, $P < 0.01$). There were no differences in

TABLE I
RESPONSE RATES FOR SELECTED VARIABLES, FIRST STUDY

		Eligible Women	% Responding
Signer	Non-specific	141	80.9
	Jewish	144	72.4
	Irish	151	73.7
Questionnaire	Long	219	74.1
	Short	217	77.1
Letter	Impersonal	215	75.9
	Personal	221	75.2
Decade of age	20	68	86.8
	30	72	75.3
	40	77	79.2
	50	77	79.2
	60	71	70.8
	70	71	62.0
Total		436	75.6

TABLE II
RESPONSE RATES FOR SELECTED VARIABLES, SECOND STUDY

		Eligible Women	% Responding
City	Newton	395	68.4
	Worcester	386	61.2
Signer	Irish	267	67.6
	Irish, professor	256	64.7
	Non-specific	258	62.1
Question on schooling present	Yes	391	62.5
	No	390	66.7
Decade of age	20	195	68.3
	30	196	68.7
	40	195	66.6
	50	195	54.8
	Total	781	64.8

response to the names signed to the letter even when the ethnic connotations of the recipients' names were reviewed. In addition, 34.1% of the women receiving a certified second follow-up and 40.7% of those in the 'telephone attempt' group responded. This difference did not achieve statistical significance but the number of observations is quite small.

Table II shows final response rates based on the 781 women deemed eligible for the second study. Controlling for 'city' in the analysis, the trend of decreasing response with increasing age is significant (summary $\chi^2 = 6.83$, $P < 0.01$). Controlling for age, the difference between the two cities approaches significance (summary $\chi^2 = 3.53$, $P \approx 0.06$). The difference between the response rates for the questionnaire with or without the schooling question is not significant overall. However, among Worcester women aged 50+, and only these women, the schooling question markedly decreased the response rate. The response rate did not vary consistently or significantly by name or title of signer of letter even after controlling for age and city.

DISCUSSION

Scott (1961) in Britain and Sletto (1940) and Sirken, Pifer, and Brown (1960) in the United States reported studies in which comparable groups received one of two short questionnaires or a long questionnaire which was a combination of the two short ones. In no report was there a significant difference in response between the long questionnaire and the two short ones. The present findings also imply that little or no adverse effect results from increasing questionnaire length. Somewhat

at variance with this is the finding of Cartwright and Ward (1968), who asked about the family planning advice offered by 232 general practitioners in Sheffield, England. They found a significantly greater response (90% v. 73%) to a short as opposed to a longer questionnaire. However, the questionnaires differed appreciably in content and precision of required responses.

There were two components to 'personalization': stamping v. franking and verbal personalization. With respect to the former, Scott (1961) found a difference in favour of the stamp (93.3% v. 89.2%), statistically significant on the large numbers in his study. Robinson and Agasim (1951) found a significant difference in favour of the stamp (73.8% v. 66.3%), but it is not clear that the study groups were comparable. They also found that the difference resulted from a poor return of franked envelopes from small towns where many of the franks were covered by a stamp. Clausen and Ford (1947), on the other hand, in a small study (no details given) reported no difference in response with stamped v. franked envelopes. There may be an advantage to stamping rather than franking but it is probably small. With respect to verbal personalization Clausen and Ford (1947) found no difference in response between veterans receiving a letter with an inside address and personal salutation and comparable veterans receiving the same letter with no inside address and the salutation 'Dear Veteran'. The present data do not allow separation of the effects of stamping from verbal personalization but they strongly suggest no appreciable difference in personal v. impersonal as used.

It seems reasonable to conclude that the 'instrument' (the form of the letter, its accoutrements and enclosures) makes little difference in response. It also seems that when a specific characteristic of the instrument is associated with differing response rates (e.g., the presence of the question on schooling) it is largely due to one subgroup of recipients (the Worcester elderly). Cartwright (1964) found that the addition of a question on schooling to a one-item questionnaire on hospitalization reduced response from 93.6% to 90.7%. It is not stated whether the lower figure was due to a particularly low response from a sub-segment of the population, e.g., the elderly.

Both the present studies show a decreasing response with increasing age. The second study suggests also that the social class of the recipient is important as Newton is a rather prosperous suburb of Boston whereas Worcester is a large city with a cross section of social classes. Neither con-

clusion is original but they have not been clearly documented in the past.

The value of the present studies lies in their placing in perspective the relative importance of the instrument and the characteristics of the target population in determining response. This is not to say that elderly or lower social class persons would always respond poorly—perhaps questionnaires concerned with matters other than health and disease would elicit a reversal of these findings. Nonetheless, it seems that money and effort spent in personalizing and 'improving' a mail instrument may not yield comparable improvement in response rate.

SUMMARY

Two studies of the determinants of response rates of Massachusetts women to postal questionnaires are described. They suggest that the characteristics of the target population are more important than the instrument in determining response rate.

We are indebted to Dr. Jacob Feldman for advice.

This work was supported by Program Project Grant 2-P01-CA06373 from the National Cancer Institute, United States Public Health Service.

REFERENCES

- CARTWRIGHT, A. (1964). *Human Relations and Hospital Care*. Routledge and Kegan Paul, London.
 —, and WARD, A. W. M. (1968). Variations in general practitioners' response to postal questionnaires. *Brit. J. prev. soc. Med.*, **22**, 199.
 CLAUSEN, J. A., and FORD, R. N. (1947). Controlling bias in mail questionnaires. *J. Amer. statist. Ass.*, **42**, 497.
 ROBINSON, R. A., and AGASIM, P. (1951). Making mail surveys more reliable. *J. Marketing*, **15**, 415.
 SCOTT, C. (1961). Research on mail surveys. *J. roy. statist. Soc., Series A*, **124**, 143.
 SIRKEN, M. G., PIFER, J. W., and BROWN, M. L. (1960). Survey procedures for supplementing mortality statistics. *Amer. J. publ. Hlth*, **50**, 1753.
 SLETTO, R. F. (1940). Pretesting of questionnaires. *Amer. sociol. Rev.*, **5**, 193.

Correction

On page 164 of the August issue of this journal equation (1) should read

$$\chi^2 = \frac{(|\sum m_1 - \sum \mu_1| - 0.5)^2}{\sum \sigma_1^2} \quad (1)$$