Quality & Quantity **26**: 233–244, 1992.

The effect of questionnaire length on survey response

BRENDAN BURCHELL
Social and Political Sciences Department, University of Cambridge, Free School Lane, Cambridge

and

CATHERINE MARSH
Faculty of Economic and Social Studies, University of Manchester, Oxford Road, Manchester

Abstract. Survey textbooks suggest that long questionnaires should be avoided, and a careful reading of the available empirical evidence confirms the negative effects of substantial length on both response rates and the quality of those responses which are obtained. Data is presented from a lengthy survey in Britain in 1987. Analysis of reasons for nonresponse to this survey suggest that length may indeed have been a significant disincentive to respond for many. However, no effect of length was found on item quality as measured by the number of responses given to open-ended questions. Unexpectedly, the variance in number of responses was greater when the questions were asked later in the questionnaire. The results are interpreted as resulting from the greater power that respondents gain as the survey proceeds.

1. The problem

Survey research is an expensive business; having gone to the expense of tracking respondents down, we often try to get as much from them as possible. While most would concede that increasing length beyond a certain point leads to an unacceptable reduction in both response rates and the quality of the data collected, there is no agreement about the timing of this critical point. In this paper, some findings are reported from a lengthy survey of work attitudes and work histories. We dwell briefly on the reasons for refusal given by non-respondents to this survey, which are commonplace but nonetheless frequently ignored by survey designers. The core of the paper concerns the results of an experiment conducted to test the effect of length of questionnaire on the responses to open-ended questions.

2. The literature

The message from most survey textbooks is clear: don't make questionnaires too long. Failure to observe this rule will produce a decline in response rates,

not simply because subjects become less willing to respond (Festinger and Katz, 1965: 49; Brook, 1977: 127) but also because the number of interviewer contacts per day decline (Backstrom and Hursh-Cesar, 1981: 252) and because evening interviewing becomes so difficult (which also leads to response bias (Sheatsley, 1983: 223)). Furthermore, it is claimed, respondents are more likely to break off the interview in the middle (Sheatsley, 1983: 223), item nonresponse increases (Anderson *et al.*, 1983), and, perhaps most importantly, responses obtained towards the end of the interview can be of lower quality because the respondent tends to become fatigued and less motivated (Goode and Hatt, 1952: 137; Festinger and Katz, 1965: 49; Cannell and Kahn, 1968: 539; Sheatsley, 1983: 223; Courtenay, 1977: 29). However, some influential voices sing a different song; most widely cited is Bradburn's (1979) suggestion than increasing the length of interview can actually enhance the importance of the study in the respondent's eyes and improve the quality of response.

The experimental literature to test these hypotheses is often weakened because the effects of length are often insufficiently disentangled from the effects of the content of the questionnaire; since content is now agreed to be about the most important factor influencing response rates (Heberlein and Baumgartner, 1978; Goyder, 1982; 1988), this is a crucial objection. However, in general, the evidence supports the prescriptions of the textbooks.

When long surveys have been compared with substantially shorter ones, they consistently (if not individually significantly) produce lower response rates (Cartwright and Ward, 1968; Blumberg *et al.*, 1974; Sharp and Frankel, 1983), although when the differences in length of questionnaire are not as great, the results show little effect (Mason *et al.*, 1961; Sheth *et al.*, 1973). The best way to study the effect of length is by comparing responses to shorter questionnaires with a longer one which is a sum of these. When response rates to the shorter versions themselves differ significantly (as in Sletto, 1940 and Scott, 1961), no conclusions can be drawn. But when they do not, it is clear that length has a detrimental effect on response (Sirken *et al.*, 1960; Berdie, 1983). We should note that Berdie himself declared his results insignificant, but on the basis of an incorrect application of the chi-square test; Kendall's tau-*b* or any of several tests of significance which make use of the ordinal information would yield a p value of less than 0.03 (two-tailed) for his results.

Non-experimental, multivariate analysis of large numbers of different postal surveys confirms that long questionnaires produce lower response rates, each additional page of the questionnaire reducing response rates by around 0.4% (Heberlein and Baumgartner, 1978; Goyder, 1982). In a similar

study of personal interviews, Yu and Cooper (1983) find a much weaker and insignificant bivariate association between length and response, but if they had controlled for sponsor (especially government *vs* private) they might have uncovered the stronger relationship found by Heberlein and Baumgartner and Goyder after similar controls. Once respondents have agreed to be interviewed, however, length appears to be only weakly (Ferber, 1966) or not at all (Sharp and Frankel, 1983) related to item non-response. The general public, when asked their opinions about survey interviewing, frequently and increasingly complain about the amount of time they take (Schleifer, 1986; Roper, 1986).

The mechanism of precisely *how* length affects response probably differs for mail and personal surveys. With mail surveys, it does appear to be the actual number of items and not the physical length of the questionnaire that puts people off (Layne and Thomspon, 1981; Champion and Sear, 1969). In personal surveys, it must presumably be the doorstep interaction between respondent and interviewer where the effects occur, but studies vary in whether refusals occur before or after the length of time is actually mentioned (Sharp and Frankel, 1983; Morton-Williams and Young, 1987).

In summary then, there appears to be a consistent negative effect of length on response rates. On the other hand, it is not very strong, and may not provide much of a disincentive to questionnaire designers who wish to maximise expensive interview time. With careful attention to the packaging of the whole survey, it may be possible to get acceptably high response rates even to long questionnaires (Dillman, 1978).

More worrying perhaps is the evidence that length produces a decline in response quality towards the end of the interview.

When the question type is open-ended, higher quality may be reflected in more answers. College students have been shown to give fewer answers to open-ended questions towards the end of a long interview (Johnson *et al.*, 1974). Furthermore, respondents appear more prone to response sets in batteries of items with similar response scales when these appear towards the end of the questionnaire, especially on subject matters of low interest to them (Herzog and Bachman, 1978). Low validity as assessed by the multimethod multitrait approach is associated with position of an item on a questionnaire. In a study of the contribution of various factors to the methods' variance in a large number of different surveys, Andrews (1984) found that position on the questionnaire accounted for between 15 and 24% of the variation in measurement quality; construct validity was low at the very beginning of a questionnaire (for around the first twenty five items), and declined once more after one hundred questions. Respondent cooperation may also decline as time goes on; Sharp and Frankel (1983) found that 27%

of people refused to be re-interviewed after a 75-min interview, compared with 13% after a 25-min one.

The evidence from the empirical literature is therefore generally consistent: length beyond around 100 questions or 10 pages of a typical schedule has a detrimental effect on both response rates and response quality.

3. The data

The survey reported below formed part of a programme of research conducted by the Social Change and Economic Life Initiative under the auspices of the British Economic and Social Research Council: sixty academics representing four disciplines from ten universities studying six different labour markets in Britain in 1986/7. Given such a large enterprise, there was inevitably a great deal of pressure to extend questionnaires to cover items of interest to everyone. The results discussed here came from a pilot survey in which an experimental attempt was made to see if length of questionnaire affected responses to open-ended questions. The schedule covered a wide variety of topics, but most of them on the theme of employment or the lack of it. As well as a large number of factual and attitudinal items, almost half an hour of the schedule was devoted to collecting systematic information on the respondents' work and life histories. The questionnaire typically took one hour and a half to complete, but the history section ensured that the variability in completion times was great.

The 300 respondents in this survey were all adults aged between 20 and 60. The sample was drawn from six large towns in England and Scotland. Different quotas were set within each of these localities, specifying the sex and employment status of each subsample of 50 respondents. Although these quotas were fixed for piloting reasons only and were not intended to give a representative sample, taken as a whole the sample is not too dissimilar to that which might have been obtained by a more sophisticated sampling procedure. For instance, the sex breakdown was exactly half males and half females; 11% of the sample were self-employed, 49% were employees, 19% were unemployed and the remaining 21% were economically inactive. However, for the purposes of analysing the experiment, it is the size rather than the composition of the sample that is more important; there were 150 in each of the experimental and control groups.

Fieldwork took place in early 1987, using interviewers employed by Public Attitude Surveys Ltd. The respondents were approached and interviewed in their own homes.

Table 1. Reasons for refusal to respond

	N	%	
"The interview is too long"	24	9	44
"No time"	48	19	
"Too busy"	24	9	
Involved in an activity	18	7	
"Not interested"	49	19	38
Too old or sick	10	4	
Principled objections to surveys	22	9	
Other	15	6	
No reasons stated	47	18	
Total	257	100	

Source: Social Change and Economic Life Initiative, 2nd pilot.

4. The effect of length on response rates

The interviewers were instructed to keep a record on their contact sheet of the number of refusals they got before each successful interview, and the reasons given at the time for the refusal. The individuals reasons were often ambiguous, and many could have been covertly referring to the time the survey would take: 'I'm getting the meal' or 'I've got an appointment with the doctor' may refer to the fact that the survey is too long given the respondent's other constraints. Table 1 shows a summary of the reasons given.

Nine per cent said explicitly that the schedule was too long, but up to 44% of respondents (61% of all refusers who gave a reason for refusal) could have been influenced by the length of the questionnaire (in one of the first four categories of Table 1). This breakdown of reasons for non-response in Table 1 seems typical for a personal interview survey of this kind (it is, for example, very similar to the profile given by Morton-Williams and Young (1987).

While it is not possible to draw any strong conclusions about the effect of length on response rates directly from this evidence without control groups with shorter schedules, the implications should certainly be a cause for concern among researchers contemplating the use of long schedules. Survey designers are regularly confronted with evidence from non-respondents that the length of their survey is an important factor, yet frequently ignore this, or even dismiss it as a rationalisation for refusal to which not much attention should be given; the prescriptive and empirical literature discussed above

almost without exception fails to comment on this most commonplace piece of survey evidence.

5. The experiment

A battery of closed questions followed by two open-ended questions were asked in one of two places in the questionnaire, either near the beginning of the questionnaire just after the life and work histories section had been completed, or near the end of the questionnaire among some items on attitudes to work. The median time for the first position was after 37 min, and for the second position after 73 min, either side of the 100 question divide which seemed important from the literature. Since the context as well as timing varies when position is rotated, we only report the effect of position on the open-ended questions; in this way we control for the effects of the immediate context, since in both positions the open-ended questions followed the same battery of closed questions.

Two open-ended questions formed the dependent variables of the experiment:

> "Why do you think unemployment is high in Britain at the moment?"

and

> "What sort of people have difficulty in finding a job?" followed by the prompt: "Why do you say that?"

Interviewers were told to probe fully for answers with phrases like "What other reasons?" and "What others?". The responses were coded on a frame which operated at two levels: broad categories (such as government blaming) and detailed explanations within these (such as personality of Mrs Thatcher). Up to four separate reasons for unemployment were coded from the first question and up to twelve responses from the second question and follow-up.

In order to verify that the allocation of respondents to the two types of questionnaire was truly random, the demographic characteristics of the two groups were compared. The results were perfectly satisfactory; the two distributions did not differ on sex, age or town. Moreover, the location of the rotated section had no effect on the overall administration time for the schedule.

6. Results of the experiment

Two hypotheses about the effect of position on open-ended questions were derived from the literature, one about the number of responses that would be obtained and the second about the validity of those responses.

Hypothesis 1. The number of responses to the open-ended questions will diminish when the questions appear later in the questionnaire (after Johnson *et al.*, 1974)

This hypothesis was tested initially by a simple *t*-test, with the total number of responses to open-ended questions as the dependent variable, and the position of the rotated section as the independent variable. This hypothesis was rejected; the means were in the opposite direction to that predicted, with more responses later in the schedule than early (5.25 and 5.01 respectively). The differences were not significant ($t_{298} = 1.03$, $p = 0.3$ two-tailed).

However, there was an unanticipated but significant difference in the variances in the number of responses given by the two samples, with the group of respondents in the later-presentation group being markedly more heterogeneous than the early-presentation group (variances 4.84 and 3.28 respectively, $F_{153,145} = 1.48$, $p = 0.018$, two-tailed). The distribution of both groups was approximately Gaussian, with no evidence of a ceiling or floor effect that could account for the difference in variances. It seems that the sitting through a lengthy survey interview genuinely makes some respondents talk more and others less.

In order to get more understanding of the factors leading to more and less reserve, the demographic characteristics of people who talk more and less in a survey were investigated. Although there were no differences in this respect between different age-groups and sexes, there was a difference in both means and variances of the number of responses between social classes: white-collar respondents had a higher mean and standard deviation (5.36 and 2.09) than blue-collar respondents (4.77 and 1.73), outcomes that were clearly significant for the means ($p = 0.023$) and borderline significant for the standard deviations ($p = 0.06$). The joint effect of class and position in the questionnaire is shown in Table 2.

The position of presentation effect seems statistically significant from the two-way analysis of variance of the results in Table 2 ($F_{1,230} = 3.82$, $p = 0.05$), with the respondents who answered the open-ended questions later in the questionnaire giving 10% more responses than the respondents who came across those questions earlier. However, since the one-way effects of position were insignificant, the results are probably an artefactual result of excluding those who had missing data on social class.

Table 2. The effect of social class and position in questionnaire on mean number of responses to open-ended questions

			Social class		
			White collar	Blue collar	All
Position	Early	$\bar{X}$	5.06	4.61	4.92
		s	1.92	1.59	1.83
		N	(80)	(36)	(116)
	Late	$\bar{X}$	5.67	4.90	5.40
		s	2.23	1.85	2.12
		N	(76)	(42)	(118)
	All	$\bar{X}$	5.36	4.77	5.16
		s	2.09	1.73	1.20
		N	(156)	(78)	(234)

Missing $N = 66$
Source: Social Change and Economic Life Initiative, 2nd pilot.

In summary, the initial hypothesis must be rejected: if anything, people seem to say more in response to an open-ended question at the end of a lengthy interview. We also discovered, unexpectedly, that the variance in the number of responses to open-ended questions increased later in the questionnaire. The interpretation of this is discussed further below.

Larger numbers of responses to open-ended questions could mean one of two different things: genuinely more opinions being proffered or greater verboseness and repetition. The first interpretation accorded better with the data: the analysis was repeated using as a dependent variable the number of different broad *categories* of response rather than the total number of responses; although the differences were smaller and less significant, they were in the same direction as before with the later presentation and higher social class being associated with higher means and variances.

The previous literature led us to expect a decline in validity in later positions. The results of testing hypothesis 1 has already cast some doubt on this, however: the increase in the content of answers given to the same questions when placed later in the questionnaire hardly suggests a decline in validity; nevertheless it was important to test the validity hypothesis directly.

Hypothesis 2. The predictive validity of responses to open-ended questions will decrease when the open-ended questions are presented later in the schedule, due to increased error (after Hertzog and Bachman, 1981 and Andrews, 1984).

This hypothesis initially stemmed from the prediction that error variance would increase after the 100 item point; since we did not expect correlated

error to be a major problem with open-ended questions, predictive validity was therefore hypothesized to decline.

Social class and voting preferences were chosen as the criterion variables, since explanations for unemployment vary sharply by these variables (e.g. Furnham, 1982). Thus, for example, while it was expected that Conservative voters would be more likely to give individualistic explanations of unemployment than Labour voters, the strength of this relationship was predicted to be greater if explanations were elicited earlier.

This hypothesis was tested by looking at the strength of association (Spearman's correlations) between each respondent's use of each of 17 broad response categories and the respondents' voting intention in three categories: Conservative, Alliance and Labour. Voting intention did have the predicted strong effect on explanations for unemployment, but, contrary to our initial prediction, there was no evidence of decline in predictive validity later in the questionnaire. In no case was there a significant difference between any of 17 pairs of correlations (using Fisher's Z test, all $ps > 0.1$, two-tailed). In fact, there was no overall pattern of stronger correlations in either early or late position; seven of the correlations were larger in the earlier position compared to 10 in the later position. The analysis using social class as the criterion variable led to the same conclusions.

Thus, while these long interviews may have contributed to general sample response biases, we have not replicated the doubts expressed in the literature about the number or quality of responses given at the end of long interviews. We have instead found an interesting result: the variation in responses to open-ended questions can increase significantly as time goes on. In the next section we suggest a mechanism whereby this might operate.

7. Interpretation: the dynamics of the survey interview

Less control is influenced in the survey interview over eliciting open-ended responses than over presenting and recording structured questions. While this does not produce generally higher response effects (Sudman and Bradburn, 1974: 60–61), effects which emanate specifically from interviewers are greater (Hanson and Marks, 1958; Shapiro, 1970; Collins, 1978; Billiet and Loosveldt, 1988). The reason for this is not hard to find; probing is hard to standardize and interviewers tend to use their own words even when standard ones are provided; indeed interviewers who are more successful in obtaining responses depart from the script more often (Morton-Williams and Young, 1986), and tend to adapt their form of words to the respondents' style (Converse and Schuman, 1974: 501).

The survey interview is a task-oriented structured conversation with a scientific purpose, and part of the interviewers' job seems to be to teach the respondent the appropriate rules of the game (Sudman and Bradburn, 1974: 4–9). However, many of the norms of daily interaction still apply, and the result is a subtle process of negotiation, compromise and accommodation between two individuals. Our results perhaps suggest the following scenario in a long interview.

At the start of the interview the respondents find themselves in a role to which they are unused, whereas the interviewer is acting out a familiar role. This manifests itself in interviewers taking the initiative in the conversation and the respondents following their cues. In this way, the interviewers influence the number of open-ended responses that they collect, eliciting what they consider to be "about the right number". As the interview progresses, though, respondents come to feel more at ease; they regain some power in the interaction and become increasingly willing to assert themselves. Feeling less constrained by the interviewers, they are freer to be more or less forthcoming depending on their level of interest, arousal and so on. This is what accounts for the greater variance in responses later in the questionnaire.

The interpretation of the effect of social class is similar. Blue-collar respondents are widely agreed to have a more "restricted verbal code" (Bernstein, 1971). This has two results: it makes them say less in answer to survey questions, but it also less obviously produces more reliance on the interviewer's cues as to how to respond.

How then should the indicative rise in the number of answers some respondents give in the later position be interpreted? The most likely explanation is that the topic is genuinely one about which respondents had a lot to say, and later on in the interview they were less held back by interviewers cues that they had given enough information already. Caution must be exercised in generalizing from these findings to other subject matters about which respondents may not become so aroused.

Clearly, if the dynamics of the survey interview are as proposed, it is not at all surprising that these results differ from those obtained by Johnson *et al.* (1974); the results reported here were obtained from face-to-face interviews of a random cross-section of adults, whereas theirs were obtained through self-completion questionnaires administered to university undergraduates in an introductory psychology class on a very different set of topics. Indeed, this very difference in experimental results highlights the importance of the social situation on responses to open-ended questions.

The lack of diminution of quality of open-ended responses with questionnaire length is reassuring for researchers using long questionnaires containing open-ended questions. Indeed, if there is no reduction of quality in open-

ended questions but a reduction for closed-ended questions, this would seem to constitute a good reason for placing open-ended questions towards the end of a questionnaire.

Notes

Submitted January 1990. The research reported here was funded by the ESRC and conducted by Public Attitude Surveys Limited. The authors would like to thank Claire Young for valuable research assistance in preparing this paper and to Colin Fraser for his helpful comments on an earlier draft.

References

Anderson, A. B., Basilevsky, A. and Hum, D. (1983). Missing data: a review of the literature. In Rossi *et al.* (eds.), *Handbook of Survey Research*. New York: Academic Press.

Andrews, F. M. (1984). Construct validity and error components of survey measures: a structural modelling approach. *Public Opinion Quarterly* 48: 409–442.

Backstrom, C. H. and Hursh-Cesar, G. (1981). *Survey Research*, 2nd edition. New York: Wiley.

Berdie, D. R. (1973). Questionnaire length and response rate. *Journal of Applied Psychology* 58: 278–280.

Bernstein, B. (1971). On the classification and framing of educational knowledge. In M. F. D. Young (ed.), *Knowledge and Control: New Directions for the Sociology of Education*. London: Collier-Macmillan.

Billiet, J. and Loosveldt, G. (1988). Interviewer training and quality of responses. *Public Opinion Quarterly* 52: 190–211.

Blumberg, H. H., Fuller, C. and Hare, A. P. (1974). Response rates in postal surveys. *Public Opinion Quarterly* 38: 113–123.

Bradburn, N. (1977). Respondent burden. US Department of Health, Education and Welfare, National Center for Health Services Research, *Health Survey Research Methods, Research Proceedings from the 2nd Biennial Conference*, Publication No (PHS) 79-3207: 49–53.

Brook, L. (1978). Postal survey procedures. In G. Hoinville and R. Jowell (eds.), *Handbook of Survey Research*, New York: Academic Press.

Cannell, C. F. and Kahn R. L. (1968). Interviewing. In G. Lindzey and E. Aronson (eds.), *The Handbook of Social Psychology*. Vol II (2nd ed). Reading, MA: Addison-Wesley.

Collins, M. (1978). Interviewer variability: the North Yorkshire experiment. *Journal of the Market Research Society* 20: 59–72.

Courtenay, G. (1978). Questionnaire construction. In G. Hoinville and R. Jowell (eds.), *Survey Research Practice*. London: Heinemann.

Dillman, D. A. (1978). *Mail and Telephone Surveys: The Total Design Method*. New York: Wiley.

Dillman, D., Christenson, J. A., Carpenter, E. H. and Brooks, R. M. (1974). Increasing mail questionnaire response: a four state comparison. *American Sociological Review* 39: 44–56.

Ferber, R. (1966). Item nonresponse in a consumer survey. *Public Opinion Quarterly* 30: 399–415.

Festinger, L. and Katz, D. (1965). *Research Methods in the Behavioral Sciences*. London: Holt, Rinehart and Winston.

Furnham, A. (1982). Explanations for unemployment in Britain. *European Journal of Social Psychology* 12: 335–352.

Gibson, D. M. and Aitkenhead, W. (1983). The elderly respondent: experiences from a large-scale survey of the aged. *Research on Ageing* 5: 283–296.

Goode, W. J. and Hatt, P. K. (1952). *Methods in Social Research*. London: McGraw-Hill.

Goyder, J. C. (1982). Further evidence on factors affecting response rates to mailed questionnaires. *American Sociological Review* 47: 550–553.

Hanson, R. H. and Marks, E. S. (1958). Influence of the interviewer on the accuracy of survey results. *Journal of the American Statistical Association* 53: 635–55.

Heberlein, T. A. and Baumgartner, R. (1978). Factors affecting response rates to mailed questionnaires: a quantitative analysis of the published literature. *American Sociological Review* 43: 447–462.

Herzog, A. R. and Bachman, J. G. (1981). Effects of questionnaire length on response quality. *Public Opinion Quarterly* 45: 49–59.

Hoinville, G. and Jowell, R. (1978). *Survey Research Practice*. London: Heinemann.

Johnson, W. R., Seiveking, N. A. and Clanton, E. S. (1974). Effects of alternative positioning of open-ended questions in multiple-choice questionnaires. *Journal of Applied Psychology* 59: 776–778.

Jones, H. J. M., Lawson, H. B. and Newman, D. (1973). Population census: some recent British developments in methodology. *Journal of the Royal Statistical Society* (A) 136: 505–538.

Kraut, A. I., Wolfson A. D. and Rothenberg, A. (1975). Some effects of position on opinion survey items. *Journal of Applied Psychology* 60: 774–776.

Layne, B. H. and Thompson, D. N. (1981). Questionnaire page length and return rate. *Journal of Social Psychology* 113: 291–292.

Mason, W. S., Dressel, R. J. and Bain, R. K. (1961). An experimental study of factors affecting response to a mail survey of beginning teachers. *Public Opinion Quarterly* 25: 296–299.

Morton-Williams, J. and Young, P. (1986). *Interviewer Strategies on the Doorstep*. Paper presented to the British Market Research Society Conference, Brighton, March.

Morton-Williams, J. and Young, P. (1987). Obtaining the survey interview. *Journal of the Market Research Society* 29: 35–54.

Roper, B. (1986). Evaluating polls with poll data. *Public Opinion Quarterly* 50: 10–16.

Rossi, P. H., Wright, J. D. and Anderson, A. B. (1983). *Handbook of Survey Research*. New York: Academic Press.

Schleifer, S. (1986). Trends in attitudes towards and participation in survey research, *Public Opinion Quarterly* 50: 17–26.

Scott, C. (1961). Research on mail surveys. *Journal of the Royal Statistical Society* 124: 143–206.

Shapiro, M. J. (1970). Discovering interviewer bias in open-ended responses. *Public Opinion Quarterly* 34: 412–415.

Sharp, L. M. and Frankel, J. (1983). Respondent burden: a test of some common assumptions. *Public Opinion Quarterly* 47: 36–53.

Sheatsley, P. B. (1983). Questionnaire construction and item writing. In P. Rossi *et al.* (eds.), *Handbook of Survey Research*. New York: Academic Press.

Sheth, J. N., Roscoe, A. M. and Talbert, T. L. (1973). Mail surveys: can segmentation and experimentation improve them? *Public Opinion Quarterly* 37: 482–483.

Sirken, M. G., Pifer, J. W. and Brown, M. L. (1960). Survey procedures for supplementing mortality statistics. *American Journal of Public Health* 50: 1753–1764.

Sletto, R. F. (1940). Pretesting of questionnaires. *American Sociological Review* 5: 193–200.

Sudman, S. and Bradburn, N. M. (1974). *Response Effects in Surveys: a Review and Synthesis*. Chicago: Aldine.

Yu, J. and Cooper, H. (1983). A quantitative review of research design effects on response rates to questionnaires. *Journal of Marketing Research* XX: 36–44.