EXPLORING THE IMPACT OF SURVEY INTRODUCTIONS

Pamela Campanelli and Patrick Sturgis, SCPR and Nick Moon, NOP Research Group
Pamela Campanelli, SCPR, 35 Northampton Square, London EC1V OAX ENGLAND

Key Words: Nonresponse, Data quality, Doorstep interaction

In July of 1995, Social and Community Planning Research (SCPR) began a programme of research funded by a grant from the UK Economic and Social Research Council (ESRC) to take an in-depth look at the role of interviewers in the survey nonresponse process. This 18 month long programme is divided into 3 sub-projects and is in cooperation with the ESRC Centre for Micro-Social Change at Essex University and the NOP Research Group. This paper describes what has been done to date for sub-project 3. This sub-project focuses on the initial doorstep interaction between interviewer and address residents. Although typically lasting less than a minute, it is this interaction which can be crucial in determining whether or not the prospective respondent will agree to cooperate.

Overall Design

Our doorstep experiment involved 32 face-to-face interviewers from two different organisations: Social and Community Planning Research (SCPR) and the NOP Research Group. The interviewers were selected to allow for geographic spread (but excluding Scotland so as to minimise travel and hotel costs) and to allow for a range of experience levels. The interviewers in the two organisations were working on two rather different types of surveys. The NOP interviewers were working on the Political Tracking survey which is a face-to-face paper and pencil interview (PAPI) survey. In the Political Tracking survey, one person per household was interviewed as pre-selected from a probability sample of the electoral register. Thus, the NOP interviewers were seeking to interview a pre-named individual. The SCPR interviewers were working on the Family Resources Survey (FRS), an extremely detailed CAPI financial survey in which all adult members of the household participate. The FRS sample is a probability sample of addresses selected from the 1996 Great Britain postcode address file. The NOP doorstep introduction data were collected in January and the SCPR doorstep introduction data were collected in March. In both cases, the doorstep experiment applied to all households in an interviewer's workload. At all households, each interviewer was instructed to complete a contact description form (see Groves and Couper, 1994a; 1994b) as soon after the doorstep conversation as possible. For a random half of the households, each interviewer was instructed to complete a contact description form and to tape-record the doorstep interaction (see Morton-Williams, 1993). The analysis to date covers three areas: 1) the effects of the different methodologies on collecting doorstep information, i.e., tape recording versus contact description form, 2) what interviewers and respondents say on the doorstep and how this is related to nonresponse, and 3) what respondents say on the doorstep and how this is related to data quality.

Data Obtained

The average response rate for the selected interviewer areas was 61[1] percent for the Political Tracking survey and 72 percent for the FRS. For the NOP Political Tracking survey 512 addresses were assigned, yielding 494 completed contact forms in total. As all calls at an address were recorded on the same contact form, this resulted in 1404 individual calls recorded on the contact forms. 256 households were to have doorstep introductions recorded and at least one useable tape was received for 207 of these. The respective figures for the FRS were 384 interviews assigned, 322 contact forms completed at eligible addresses and 1148 calls recorded on the contact forms in total. 192 tapes were assigned, yielding 146 useable tapes. In both cases, the taping shortfall was due to several factors: 1) interviewers occasionally forgot to use the tape recorder or had problems with the tape recorder; 2) a few tapes were of such poor sound quality that they could not be analysed; 3) two interviewers from NOP had to leave the project early (for personal reasons) and were unable to complete their taping assignments; and 4) some respondents refused to give us permission to keep the tapes (this happened in 5 percent of cases for the Political Tracking survey and 4 percent of cases for the FRS).

For households in the tape-record condition, interviewers were instructed to record all calls on the household until an interview was achieved. This resulted in 401 individual taped calls for the Political Tracking survey and 447 individual calls for the FRS. Each of these was coded onto a form which could easily be compared to the original contact description form. The coding of the Political Tracking tapes was conducted by two SCPR researchers. At the level of individual codes, the reliability between coders for interviewer behaviour was .81 (representing the proportion of agreement between the two coders). The respective figure for respondent behaviour was .67. One of the two SCPR researcher/coders then coded all

[1] The overall response rate was 67 per cent for Political Tracking and 69 per cent for the FRS.

of the FRS data. In addition to the doorstep conversation, the contact description forms included items such as, outcome of call, situational and environmental factors affecting participation (many of which had been used by Groves and Couper (1995) on their forms.

Method of Collecting Doorstep Material

It was clear from our study of both modes of data collection (tape versus contact description form), that some pieces of information are best picked up by a particular method. Whereas the tape is best for studying the actual conversation elements, the contact form is best for studying all of the situational and environmental aspects of the interaction. Although there is always a concern that the use of a measuring instrument such as a tape recorder may affect the measurements being taken, there was no sign of a taping effect in the current study. There was essentially no difference in the types of interviewer and respondent comments (conversation elements) that were recorded by interviewers on the contact forms under the taped versus not taped conditions across both surveys.

Reliability between tape and contact form data

The project was designed to allow us to examine the reliability between tape and contact form methods of data collection. The reliability figures for this comparison are shown in Table 1 and are based on an analysis at the level of the individual call, prior to the re-coding of the 'Other specify' category.

TABLE 1: RELIABILITY BETWEEN DOORSTEP INFORMATION COLLECTED VIA TAPE AND VIA CONTACT FORM

	Interviewer Statements		Respondent Statements	
	FRS	NOP	FRS	NOP
proportion agreement over all codes	.48	.41	.45	.31
proportion agreement using 4 category breakdown	NA	NA	.68	.61

Overall, reliability figures suggest that the use of the contact form data should be treated with some caution. The reliability figures are much lower than would be desired. Our first suspicion is that relying on interviewer memory may reduce accuracy in the completion of the contact forms. It is possible that interviewers completing the form are influenced by the outcome of the call or content of the interview. For example, if a respondent was not very interested in the interview, would the interviewer's memory be influenced to say that he/she had made statements about not being interested? Cognitive psychology suggests that our memory for gist is far better than our verbatim memory. It also suggests that we encode and store our interpretation of a text rather than the text itself. Not only do we lose verbatim details, we also seem to add in (and recall) plausible details which fit our interpretation (Loftus and Palmer, 1974). Thus it is possible that contact description form data is prone to outcome effects and memory bias.

Conversation Elements and Nonresponse

Analysis of the interviewer and respondent conversation started with the creation of a separate dummy variable for each conversation element. This was necessary, because the original interviewer and respondent items were 'Code All That Apply' items. Dummy variables were also created to cover frequently mentioned conversation elements from the 'Other specify' category. Each dummy variable was summed over all mentions within a call. Due to the small size of the dataset, we then made the decision to sum these over all calls. Our next step was to combine what we considered to be equivalent conversation elements and to see how these new categories were related to the final outcome of the interview (see Table 2). In their work, Groves and Couper (1994a) used a four category scheme for respondent verbal behaviour: negative statements, time-delay statements, positive statements, and questions. They have also used a three category scheme: negative statements, time-delay statements and questions (see Groves and Couper, 1994b). In addition, to our own collapsing of categories, we have replicated their scheme, as shown at the bottom of Table 2.

The first part of Table 2 shows what interviewers said on the doorstep. In the full version of our paper, we explore how these statements vary by organisation, survey topic, survey design and method of collecting the doorstep material. The second part of Table 2 shows that there is a wide range of respondent behaviour between the clear opposites of 'come on in' and 'I don't want to'. Respondent statements are similar to the results from Couper (1995) who found that over all contacts, the most frequent respondent comments were 'too busy' (30 percent), 'not interested' (20 percent), and 'bad timing' (6 percent). This can be compared to the first three rows of the respondent portion of Table 2 by equating our 'presently occupied' category with Couper's 'bad timing' category. Table 2 also shows that for our data some of the most frequently mentioned comments come under the heading of Positive/Neutral statements. This was not the case in Couper's study. This can be explained by the fact that the contact description form that was used for the analysis in Couper (1995) did not capture

positive and neutral comments made by respondents other than the 'I enjoy doing surveys' category.

Table 2 also indicates (with an asterisk) which of the conversation elements are related to nonresponse. In a few cases, the initial presentation of the interviewer at the doorstep generally has a positive relationship with response. In addition, in at least one instance, having the interviewer state that he/she will call back is also positively related to response. This, of course, is a very crude analysis. It does not by any means indicate which statements are the most effective at gaining cooperation. It would also be ill advised to try to assign a causal relationship as one can not easily separate out which statements are independently volunteered by the interviewer and which are reactions to the respondent.

In terms of respondent behaviour, it can be seen, that in general, the negative statements are negatively related to response, the positive and neutral statements and the majority of questions are positively related to response. These findings support the work of Groves and Couper (1994b). In addition to these general trends, specific aspects of the table are worth examining. For example, note the positive association of overall cooperation with respondents saying they had received the advance letter versus the negative association with respondents saying they had not received it.

The Doorstep Conversation and Data Quality

Our focus thus far has been on relating what the interviewer and respondent say on the doorstep to whether or not the interviewer is successful in obtaining

TABLE 2: MOST COMMON INTERVIEWER AND RESPONDENT STATEMENTS OVER ALL MENTIONS AND ALL CALLS ⌘

	FRS		Political Tracking	
Interviewer Comments	Tape % n=146	Contact % n=322	Tape % n=207	Contact % n=494
introduced myself	75	90 * +	42	61 * +
said where from	85	88	93	80 * +
showed/handed over id card	27	87 *+	41	70 * +
mentioned the survey/details of survey	83	75	94	63 * +
mentioned/handed over advance letter	81	89 *+	1	2
mentioned/handed over leaflet	3	13	0	0
asked if correct address/respondent	49	57 *+	7	5
asked how many adults at address	30 * -	70	0	0
need to see everyone in household	12	0	0	0
said I'll call back	49	44	56 * +	27
described how selected	27	34	66	49 *+
represents other people	10 * -	13	3	6
said interviewing others in area	8	9	4 * -	10
particularly want your views	5 * -	0	6	0
said called before	25	3	8	1
asked if convenient now	18	0	0	0
asked will you do it/may I come in	15	13 * +	19 * +	12 *+
try to arrange appointment	36	3 * +	12	2
apologising for inconvenient call	10	1	0	0
it will take about x mins; wont take long	10	1	21 * +	0
stressing confidentiality	12	6 * -	0	0
explained results could affect them	4	17	0	2
said topic should be interesting/enjoyable	3	12	6 * -	16
topic is important	5	2	1	0
described use of computer	1	5	0	0
said not a salesperson	0	1	5 * -	4
miscellaneous friendly	12	6 * +	0	0
other	19 *-	14 *-	23	23

Table 2 Continued	FRS		Political Tracking	
Respondent Comments	Tape % n=146	Contact % n=322	Tape % n=207	Contact % n=494
Negative Statements				
not interested (in general)[2]	8 * -	17 * -	12 * -	16 * -
too busy	21	13 * -	21 * -	17
presently occupied	14 * -	16	1	6
no I don't want to	13 * -	11 * -	6 * -	4 * -
didn't receive your letter	16 * -	16	0	0
fears concerning confidentiality of surveys	5	12 * -	1	1
don't know anything	3 * -	4	0	5
waste of time/money	1	5 * -	1	2
call back (negative)	0	5	0	9
other (negative)	21 *-	21 *-	6 * -	12 * -
Positive/Neutral Statements				
come on in	56 * +	68 * +	34 * +	40 *+
have received your letter	34 * +	38 * +	0	0
call back (positive/neutral)	14	25 * +	22	9
he/she isn't in	16	12	18	11 *+
suggested good time/appointment	25	7 * +	7	5
describing patterns of being in/out	16	3	5	1
confirming correct address	10	2	0	0
saying who lives at address	12	1	0	0
that's me	2	0	5	0
they've moved out	0	0	7	4
I'll think about it	0	2	0	14 *+
other (positive/neutral)	15	9	13	21
Questions				
what's the purpose/what's it all about/who are you/who is the sponsor	12	39	7	11
how long will it take	16	19 * +	16 * +	23 *+
why/how was I chosen	2	13 * +	5	8 *+
is it compulsory	6 * -	4	0	0
other (question)	10	10 * +	6	10 * +
Groves and Couper Scheme				
All positive/neutrals	85 * +	82 * +	66 * +	58 * +
All questions	49	52	44	44 * +
All negatives	41 * -	48 * -	29 * -	32 * -
All time-delay	21	13 * -	21 * -	17

* $p < .05$; + indicates positive relationship with response; - negative relationship.
⌘ Numbers do not add to 100 percent because multiple responses are possible

[2] NOP data combines 'not interested in politics' and 'not interested (in general) in this category.

an interview. Our design, however, allows us to look beyond the impact of doorstep behaviour on nonresponse in order to examine its effect on the data obtained. Using a contact description form and data from a political attitudes survey, Couper (1995) found that what respondents said at the time the interviewer introduced the survey was significantly related to the attitudinal data they subsequently provided in the interview proper. This finding suggests an answer to an old question: Is what respondents say on the doorstep literally true or is it merely reflective of social conventions and politeness in social interaction? For example do

respondents who say they are not interested in a survey topic actually show lower levels of interest as measured by their subsequent interview data or is this simply one of many verbal techniques for declining the survey request without appearing rude and unhelpful?

Couper's finding is important from a practical as well as a theoretical viewpoint because it provides further evidence that those who respond to surveys differ in important ways from those who do not respond and that this difference is reflected in the data obtained. Our study is, in part, a small-scale replication of Couper's work.

Indicators of 'Data Quality'

Several indicators of data quality were used for the Political Tracking dataset. Cumulative measures were built to indicate the total number of occurrences of 'don't know', and 'not answered' across the whole questionnaire and of 'no opinion' on one particular political scale. We also looked at the relationship between conversation elements and two questions about willingness for future contacts: one for other surveys on the political situation and the other asking for a telephone number so that supervisors could do quality control back checks on interviewers. Also considered was an indicator of response inconsistency, based on a short 3 question scale and the respondent's substantive responses to several political questions. Similar indicators were constructed for the FRS dataset. These included the total number of occurrences of 'don't know' and 'not answered' across the whole questionnaire, the willingness of the respondent to receive a survey back check, whether the interviewer was able to check the respondent's payslip for the income question and whether the respondent was willing to divulge their total savings.

Replicating Couper (1995)

In terms of replicating Couper (1995), we were mainly interested in the relationship between respondents' comments about being 'too busy' and 'not interested' and our various indicators of data quality from the interview. Prior to this work, we looked at the relationship between these two comments and various back-ground characteristics to see if these needed to be taken into account in the analyses. The propensity to mention these statements was evenly spread across the population. The only exception among the many items examined, was that skilled manual workers were more likely to say 'not interested'. This is similar to Couper's (1995) finding that it is those individuals with less education that are more likely to say 'not interested'.

The clearest indicators of a replication of Couper's (1995) work are found in the Political Tracking contact form data. Similar to Couper's findings, the 'not interested' category is related to the number of 'don't know' answers that the respondent gave in the interview, refusal to provide income details, and the respondents' lack of consent to be contacted again. In addition, there is some indication that these 'not interested' respondents are truly uninterested in politics. For example, they were more likely to answer 'none' to the first question on the survey about political party membership. Similarly, if attached to a party, this attachment was not very strong. In contrast, no relationship was observed between these variables and the category, 'too busy'. It is difficult to discern whether the same pattern is supported by the tape data because of the smaller overall sample size. Furthermore, in the taped version of the political tracking introductions only 20 respondents who said 'too busy' and only 4 of the respondents who said 'not interested' actually went on to participate in the survey. This is because the majority of respondents who did mention that they were 'not interested' were so 'not interested' that they failed to participate in the survey in the first place. Despite this, significant associations were observed in the tape data between saying 'not interested' and the total 'don't knows' item and for a political party preference of none. Several other items appeared to show a similar trend to the contact form data (such as those relating to respondent willingness to be re-contacted in the future) but failed to reach statistical significance. For the FRS contact form data, the evidence for any kind of data quality effect was sparse. Although there were trends in a direction consistent with Couper's (1995) findings (e.g. the total number of 'don't knows' and the number refusing the total savings question were highest for those who were 'not interested'; the 'not interested' group had lower values on showing their payslips to the interviewer) only the association between saying 'not interested' and being less likely to agree to a survey back-check reached statistical significance.

Comparison with the FRS tape data is extremely difficult, again because of the overall smaller sample size and in particular because of the small number of 'not interested' cases who eventually participated (i.e., 3). Although interpreting such relationships should probably be avoided, it is interesting to note that the trend between saying `not interested' on the doorstep and refusing to answer the savings item or to agree to a survey back-check still applies.

Conclusions

Our work has been divided into three separate areas: 1) the effects of the different collection methodologies (i.e., tape recording versus contact

description form) on the resulting doorstep information, 2) what interviewers and respondents say on the doorstep and how this is related to nonresponse, and 3) what respondents say on the doorstep and how this is related to data quality.

The findings suggest that there are distinct advantages and disadvantages to each of the two methods of data collection. For example, non-verbal aspects (e.g. whether the interaction took place through a half open door or an intercom, whether the respondent was engaged in other activities when the interviewer called etc.) are best collected with a contact description form. In contrast, the tape recording is best for picking up the actual verbal interaction between respondent and interviewer. The tape recordings provide a richer fabric of verbal behaviour than could ever be captured on a contact form and with greater accuracy. On the other hand, there is always a concern in experimental research that the presence of a tape recorder could alter the behaviour of interviewers and/or respondents. Find- dings from the current study suggest that the presence of a tape recorder has not had such an effect.

Our study design allowed us to test the reliability between the tape and contact description form methods of data collection. Our data shows that the percent agreement between the two methods in terms of the verbal behaviours identified, is rather low. This is in line with the work of McCrossan (1993), who found that the contact description form data had a low reliability. Thus, the contact form data should perhaps be viewed with some caution.

The data from this study show a wide range of both interviewer and respondent verbal behaviour on the doorstep and these are similar to the findings of Morton-Williams (1993); Groves and Couper (1994a; 1994b), and Couper (1995). Although several of the interviewer statements on the doorstep were related to the success or failure to gain an interview, no simple causal relationship should be inferred. One cannot determine in which direction causality is working or whether more important, unobserved variables are mediating the associations.

The analysis of respondent behaviour is a different case. Here the findings clearly support the work of Groves and Couper (1994b). For example, the negative statements respondents make are actually negatively related to response, and the positive and neutral statements and the majority of questions are positively related to response.

In our study we also replicated Couper's (1995) work. Couper examined how respondents' statements of 'too busy' and 'not interested' on the doorstep were related to interview data quality. It should be noted that for both the Political Tracking and FRS studies these comments were less frequently used than in Couper's study. This smaller frequency has strong implications for the analysis. Generally the clearest replication of Couper's pattern of results was found with the contact form data of the Political Tracking survey. There is also some evidence for a similar pattern amongst the taped data for the Political Tracking survey. On the other hand, there was little evidence of a replication with the FRS data. This makes sense as Couper's original work was on a poli-tical attitudes survey. The difference between the data quality results for the contact form and tape data for the Political Tracking survey could be due to either the smaller sample size of the tape-coded data or the low validity of the contact form data - the associations being the result of interviewers reporting their impressions of reluctant and disinterested respondents and reporting that they said 'not interested' when in fact they may not have explicitly said this.

Our further investigations suggest there is evidence supporting both of these factors. The small frequency of respondent comments and overall small sample size are clearly distorting the results. Similarly, evidence suggests that interviewers do not remember to record potentially difficult interactions as well as they do other interactions.

References

Couper, M. (1995, April). "The impact of survey introductions on data quality", paper presented at the International Conference on Survey Measurement and Process Quality, Bristol, UK.

Groves, R., and Couper, M. (1994a). "Householders and interviewers: The anatomy of pre-interview interactions", unpublished paper, Ann Arbor, MI: Survey Research Centre, University of Michigan.

Groves, R., and Couper, M. (1994b). "Contact-level influences on survey participation", revision of paper originally presented at the annual meeting of the American Association for Public Opinion Research, Danvers, MA, May 1994.

Groves, R., and Couper, M. (1995). Personal communication.

Loftus, E., and Palmer, J. (1974). "Reconstruction of automobile destruction: An example of the interaction between language and memory", Journal of Verbal Learning and Verbal Behaviour, 13, 585-589.

McCrossan, L. (1993). "Respondent-Interviewer Interactions in Survey Introductions", paper presented at the 4th International Conference on Household Survey Nonresponse Bath, UK.

Morton-Williams, J. (1993). Interviewer Approaches. Aldershot: Dartmouth Publishing Company Limited.