

Comparing Telephone and Face-to-Face Interviewing in the United Kingdom

W.M. SYKES AND M. COLLINS¹

ABSTRACT

This paper presents results from methodological experiments comparing telephone and face-to-face interviewing in surveys of the general population. The relatively low level of telephone ownership in the United Kingdom, especially among the less privileged, argues the need for a dual-mode approach combining telephone interviews with face-to-face interviews for those without telephones. This approach depends on the absence of differential mode-effects on the answers obtained or on the ability to account for these effects when they occur.

KEY WORDS: Telephone interviewing; Dual-mode interviewing; Social surveys; Response rates; Data quality.

1. INTRODUCTION

The choice of a mode of data collection for a survey depends upon the availability of facts about the alternatives. In the U.K., such facts about telephone interviewing have just recently begun to emerge. The necessary comparisons between telephone interviewing and other data collection modes have been carried out only in the last two years. This delay is surprising given the lively debate about the merits and drawbacks of telephone interviewing and the attention which the issue has received in other countries.

Two studies conducted by the Survey Methods Centre at Social and Community Planning Research comparing telephone and face-to-face interviewing provide the focus for this paper. Carried out in 1983 and 1984, these studies examine some of the central issues: the public's willingness to take part in telephone surveys and the kind, quality and volume of data that can be collected. The studies are described in Section 2 and their results presented in Sections 3 and 4. Reference is also made to another British study – an experiment carried out in 1985 by the Market Research Development Fund – and to the larger volume of methodological research conducted in other countries, particularly the United States.

2. THE SCPR STUDIES

Our research program reflected telephone ownership which is low by North American standards: about 75% of households possessed telephones in 1983. Non-coverage is substantial and crucial, for social researchers, because of its bias towards less affluent sectors of British society. In this context, the main objective was to evaluate dual-mode interviewing, where telephone owners would be interviewed by telephone, and non-owners face-to-face.

The first study provided two comparisons towards this evaluation: between an experimental dual-mode sample and a larger national sample interviewed face-to-face; and between two samples of telephone owners, one sample interviewed by telephone, the other interviewed face-to-face. In this paper, we focus on the latter comparison, which addresses the question

¹ W.M. Sykes and M. Collins, Survey Methods Centre, Social and Community Planning Research, 35 Northampton Square, London EC1V 0AX, England

that lies at the heart of any evaluation of the dual-mode approach: are telephone and face-to-face data compatible or are there modal differences between them? If there are modal differences, the data cannot be “added” together and treated as a single data set without the kind of adjustments not usually possible in a one-time survey. The second study concentrated only on this direct comparison between the two interview methods among telephone owners.

2.1 Study 1

The first study was conducted alongside the 1983 British Social Attitudes Survey, which is here referred to as the “main” survey. This survey involved face-to-face interviews of about an hour, covering a wide range of political, economic, social and moral issues.

The sample for the main survey was about 1,750, and was representative of adults aged 18 or over living in private households. For practical reasons, the sample was confined to those at addresses in the Electoral Register. People living in institutions (though not private households at such institutions) were excluded, as were the 4% of adults known to live at addresses not on the Electoral Register (Todd and Butcher 1982).

A multi-stage design was used with four stages of selection: 103 constituencies in England and Wales and 11 local authority districts in Scotland were selected with probability proportional to electorate; within each a single polling district was selected, again with probability proportional to electorate; from each polling district, 23 addresses were selected with probability proportional to the number of electors registered at the address. At the final stage, one person at each address was selected by the interviewer, using an adaptation of the Marchant-Blyth procedure (Blyth and Marchant 1973).

For the experiment, a parallel sample of about 800 addresses (seven per area) was selected from the same 114 sampling points. These addresses, together with all the names in the Electoral Register, were submitted to British Telecom’s telephone number-retrieval facility. The facility yielded telephone numbers for 65% of the submitted addresses. Most of the difference between this retrieval rate and the level of telephone ownership – around 75% at the time – can be explained by ex-directory numbers: about 12% of telephone numbers in Great Britain are ex-directory, with regional and other variations as noted by Collins and Sykes (1987). Other problems in tracing telephone numbers seem to have had little effect.

The following procedure was used by British Telecom for retrieving telephone numbers: once the correct telephone exchange area had been identified by the address, the subscriber’s name was looked up in the directory. Specific address details (i.e., the street name) helped distinguish between subscribers with identical names. Since it is not clear from the Electoral Register which of the names at an address is that of the subscriber, British Telecom was asked to check every name before abandoning a search.

The telephone numbers obtained were systematically assigned to four sub-samples. Two of these were interviewed by telephone using a questionnaire expected to take about 20 minutes to complete. The questions were drawn from all sections of the main Social Attitudes questionnaire. The other two sub-samples were interviewed by telephone using a longer questionnaire – estimated at 40 minutes – that was also drawn from the main survey questionnaire. Sub-samples allocated to both the 20-minute and the 40-minute questionnaires were sent a letter before the telephone calls. The other sub-samples received no advance warning of the survey. In all cases the selection of a respondent for interview was on the same basis as for the main survey.

Experimental sample addresses for which no telephone numbers could be obtained from British Telecom were given face-to-face 20-minute interviews. Combined with those obtained by telephone, these interviews formed a dual-mode survey that was compared with the main face-to-face interview survey (Sykes and Hoinville 1985).

A more direct examination of interview mode effects was sought by submitting a systematic sub-sample of 600 of the main sample addresses (five in each area) to British Telecom's number-retrieval service. In this case, numbers were returned for 55% of the addresses (the variability of the success rate of the British Telecom number-retrieval service remains unexplained). Comparisons were then made between those who were interviewed by telephone and those who could have been interviewed by telephone but were interviewed face-to-face. By restricting comparisons to the telephone-accessible population, we controlled for effects attributable to differences between the compared populations rather than to differences in the mode of data collection.

2.2 Study 2

The second experiment concentrated on this direct comparison. About 2,300 addresses were selected from the Electoral Register, as in Study 1, and were sent to British Telecom for telephone numbers (with in this case, a 61% retrieval rate). Addresses for which telephone numbers were retrieved were split into three sub-samples. One group was interviewed by telephone using "pencil and paper" methods; another was interviewed using Computer Assisted Telephone Interviewing (CATI); the third was interviewed face-to-face. Our experiment with CATI was a practical failure (for a number of reasons), but the other two sub-samples again give us a direct comparison between people interviewed by telephone and people who could have been interviewed by telephone but were interviewed face-to-face. The questionnaire, designed to take 25 minutes, consisted of a sub-set of questions from the 1983 British Social Attitudes Survey.

2.3 Limitations on the Comparisons between Interviewing Modes

Three factors could limit comparisons between the answers obtained face-to-face and those obtained over the telephone. First, differential non-response (as discussed in Section 3) could have led to differences in the composition of the respondent sets. This possibility was tested using a number of demographic and socio-economic variables believed to be associated with certain attitude variables. Significant differences between the respondent sets suggest that, quite apart from any differences between the modes in overall response levels, certain kinds of people are more likely to participate in a telephone rather than a face-to-face survey, and vice versa. The variables examined were: age within sex, marital status, household composition, economic status, socio-economic group and geographical location. No statistically significant evidence of differential non-response was found in the first study. In the second study, two variables showed statistically significant differences between the telephone and face-to-face samples: household composition (the telephone respondents included a higher proportion of childless couples under 60, while the face-to-face sample had a higher percentage of couples with young children and teenagers); and socio-economic group (intermediate and junior non-manual workers and those in "other" occupations had greater representation in the telephone sample than face-to-face, and "homemakers" were a higher proportion of the face-to-face sample). These differences may well represent only sampling fluctuations, but they should lead to some caution in the interpretation of differences between the answers of the two samples.

The second possibility is of different levels of skill or supervision between the telephone and face-to-face interviewers. Six telephone interviewers were employed on the first experimental survey. Two were fully trained and experienced face-to-face interviewers, but the remainder had had no previous interviewing experience and so received basic interviewer training as well as the special telephone interviewing training that all six interviewers underwent. The second study involved 10 interviewers, three of whom had worked on the previous study. As in the previous study, a supervisor was present to listen in, advise on interviewing technique when necessary and check for obvious errors in completed questionnaires.

The face-to-face interviewers for both studies were drawn from Social and Community Planning Research's panel of about 300 regularly employed face-to-face interviewers. Their training in basic interviewing techniques was similar to that given to the telephone interviewers. However, for the most part, the face-to-face interviewers were more experienced than their telephone counterparts. Differences between the two groups of interviewers should, therefore, be kept in mind, especially differences suggesting lower quality in the telephone interviews.

The third factor is the questionnaires. The main Social Attitudes questionnaire, comprising about 100 questions, was divided into five broad topic areas: employment, education, health and housing, issues of social class, and racial and sexual equality. The experimental questionnaires were composed of those questions considered most important in the main survey. These questions were chosen to represent the full range of question types in the main questionnaire.

As a result, the experimental questionnaires covered a range of topics (including some "sensitive" issues) and included questions involving different kinds of response tasks and levels of complexity. The order of the questions on the Social Attitudes Survey was maintained for both the 20-minute and 40-minute experimental questionnaires used in the first study and for the 25-minute questionnaire used in the second study. Thus the 40-minute questionnaire was not made up of the short questionnaire followed by a further 20 minutes of questions: rather, questions from the 20-minute version were spread throughout. Alterations to question wording were made only when unavoidable; for example, re-wording to adjust for the necessary absence of showcards. The Social Attitudes Survey questionnaire consists largely of closed questions, so few of the results from our experiments relate to open questions.

All of these limitations should be kept in mind when examining our results, but they are largely inevitable in such comparative studies. As described above, we have tried to identify and minimize them. They are of great concern only when our results suggest mode effects that might confound the effects of other variables: most of our results do not point to this. Thus the limitations should be considered only as potential sources of effects counteracting mode effects we might otherwise have found – surely a less serious threat to the validity of our conclusions.

3. RESPONSE RATES

In the U.K., doubts about the feasibility of telephone interviewing, particularly for social surveys, stem from concerns not only with the level of communication possible, and its effect on both cognitive and affective dimensions of the interview, but also with the general social acceptability of this use of the telephone. In Britain, it is a common belief among researchers that "cold calls" from strangers are likely to be treated with circumspection: a call from a telephone interviewer may be regarded as inappropriate and intrusive.

A common counter argument points out the possible advantages telephone interviewing has over face-to-face interviewing, particularly in inner city areas. Escalating personal and property crime has led to increasing suspicion of strangers, which means falling response rates and the installation of devices such as entry-phones that make it harder for personal interviewers to contact respondents. By telephone, contact will also certainly be made at an address if someone is there, and, if not, subsequent attempts are not expensive.

Table 1 shows the response rates for both studies conducted by the Survey Methods Centre.

Table 1
SCPR Experiments: Response Rates

Bases	Study 1		Study 2	
	Telephone (429)	Face-to-Face (313)	Telephone (730)	Face-to-Face (631)
	%	%	%	%
Completed interviews	53	60	46	68
Partial interviews	1	-	-	-
Refusal (no selection)	5	2	21	6
Refusal (proxy)	9	5	7	4
Refusal (selected person)	11	18	10	11
No contact ^a	3	1	8	4
Selected person never in	3	3	3	2
Ill, away, language problems	2	5	2	4
Other ^b	13	6	4	2

Study 1: $\chi^2 = 3.72$ d.o.f. = 1 $0.05 < p < 0.1$

Study 2: $\chi^2 = 66.22$ d.o.f. = 1 $p < 0.001$

Studies 1 and 2 combined: $\chi^2 = 59.46$ d.o.f. = 1 $p < 0.005$

} comparisons with only two categories:
completed interviews and
non-completed interviews.

^a Includes "Ring no answer" and "Permanently engaged".

^b Includes "Broken appointments", "Too old", "Incapacitated", "No connection", "Right number, wrong address".

Response to these studies, for both the telephone and face-to-face components, was relatively low. (We would normally expect personal interview response rates of over 70% before reissue of refusals.) This owes something to the nature of the surveys – general purpose surveys are notoriously difficult to "sell" to respondents. The same argument can also be applied to the only other major British methodological comparison survey, carried out by Marplan on behalf of the Market Research Development Fund. This study used the same sampling method as our own experiments and also included a wide range of general questions, under the title *Lifestyle in the 1980's*. In this case, the response rates obtained were 45% by telephone with a sample base of 1697 and 67% face-to-face with a sample base of 1233 (Market Research Development Fund 1985). In both our studies, the response rate was lower for telephone interviews: barely half of the issued addresses yielded interviews. As Table 1 shows, the difference was on the borderline of non-significance for Study 1 but was statistically significant for Study 2 and for Studies 1 and 2 in combination.

The difference might be attributed to our relative lack of experience with telephone interviewing, but it is consistent with findings from other countries. For example, in the United States lower response rates – mostly arising from the higher incidence of refusals to cooperate – have been reported by a number of authors (e.g., Hochstim 1967; Henson, Roth and Cannell 1977). The position is summarized by Groves and Kahn, who write:

"The response rate of national surveys remains at least five percentage points lower than that expected in personal interview. This has been a rather stable comparison despite changes over time in training of interviewers, monitoring techniques, feedback procedures from monitors, and techniques of introducing the survey to the respondent." (Groves and Kahn 1979; p. 219)

These findings suggest that sociological and psychological explanations of resistance to the telephone approach may be more appropriate than explanations of interviewer and general

methodological inexperience. However, the first SCPR study appears to have been rather more successful than either the second or the MRDF study. It has been suggested that this difference was due to the interest and excitement surrounding the first experiment. This may have communicated itself to the interviewers (for example, researchers were continually “dropping in” to observe the proceedings), thus affecting their success rates. Certainly, experience with face-to-face surveys suggests that interviewer morale and energy are important for good response rates.

In the SCPR studies two survey conditions were varied to assess their impact on telephone response rates. For the first survey, half the telephone respondents were asked to do 20-minute interviews and the other half did 40-minute interviews (respondents were told the length of the interview towards the end of the introduction), and in both surveys advance letters giving notice of the interview were sent to a random half of the telephone sample.

Table 2 shows that response for the 40-minute interview was lower than for the 20-minute interview, although the difference between the overall distributions was not significant. The main single reason for this lower response was the higher direct refusal rate, possibly indicating that respondents were less willing to undertake the longer interviews. However, very few respondents who had agreed to participate terminated an interview prematurely – even with the longer interview.

Different strategies may be needed for longer questionnaires. While it may be reasonable to request respondents to take part in a 20-minute interview at the time when first contact is made, a system of appointments may be more successful where more interviewing time is required. Wiseman and McDonald (1979) suggest that refusal rates are likely to be lower when interviewers are instructed to make call-back appointments should the respondents indicate that they are busy.

In other studies, sending advance letters to potential telephone respondents has been found to improve response rates. For example, Dillman, Gallegos and Frey (1976) obtained refusal rates which were, on average, 6% lower for respondents receiving advance letters (compared with 14%). As Table 3 shows, in the SCPR experiments response rates were slightly higher among respondents who had been sent an advance letter (no record was kept of whether letters had been received) although the differences were not statistically significant.

To explore why respondents refuse to be interviewed by telephone, 55 refusers to the first study were followed-up to see whether they would have co-operated at the first contact if they had been approached personally. Forty said that the method of interview would have made no difference to their decision, and only a very small number of these people subsequently agreed to be interviewed. Most of the rest said they would have taken part if they had been approached face-to-face and eventually completed a face-to-face interview (13 out of 15).

Because face-to-face refusers were not followed up, we do not know if a proportion of this group would have preferred to be approached by telephone.

3.1 Response Differences and Data Quality

The public's perception of the proper use of the household telephone may effect not only response rates, but also the kinds of questions respondents will be prepared to answer. Of even greater concern, however, is the type of communication possible between interviewer and respondent and its potential effect on the measurements made.

Face-to-face communication takes place both verbally and non-verbally, while the telephone has only limited channel capacity with exchanges between interviewer and respondent restricted to what is said and so-called paralinguistic cues: tone of voice, pauses and so on (Miller and Cannell 1982).

Table 2
SCPR Experiments: Effects of Interview Length (Study 1)

Bases	40-Minute (206)	20-Minute (223)
	%	%
Completed interviews	48	59
Refusal	27	23
Other	25	18

$$\chi^2 = 4.7 \text{ d.o.f.} = 2 \quad 0.10 > p > 0.05$$

Table 3
SCPR Experiments: Effects of Advance Letters on Response Rates

Bases	Study 1		Study 2	
	Letter (215)	No Letter (214)	Letter (388)	No Letter (392)
	%	%	%	%
Completed interviews	55	51	48	43
Refusal	23	27	37	38
Other	22	21	15	19

$$\text{Study 1: } \chi^2 = 1.09 \text{ d.o.f.} = 2 \quad p > 0.5$$

$$\text{Study 2: } \chi^2 = 2.8 \text{ d.o.f.} = 2 \quad p > 0.2$$

$$\text{Studies 1 and 2 combined: } \chi^2 = 3.49 \text{ d.o.f.} = 2 \quad p > 0.1$$

The possible implications for survey measurements of the telephone's limited channel capacity are numerous. For example, the absence of visual aids may increase the difficulty of some response tasks. "Voice only" communication may not convey the full meaning behind respondents' words (making it difficult, for example, to probe open-ended questions) and may not reveal if they actually understand the questions. There may also be limitations on the interviewer's ability to perform his or her role. Can verbal signals, for example, replace the non-verbal cues that convey interest and attention to the respondent, or those that help control the interview? Can the interviewer hold the concentration of the respondent, particularly in long interviews? Conversely, is the absence of visual stimuli a desirable reduction in the many sources of variability in survey data? Finally, does the greater social distance in the telephone interview make the respondent more or less comfortable in revealing sensitive information such as income, or information with a strong social desirability component?

SCPR's experiments addressed some of these issues.

3.1.1 General Comparisons

Given the different refusal rates of the interviewing modes, it is surprising that there are few other general differences. This result has been replicated in many studies in the U.S. (Groves and Kahn 1979; Lucas and Adams 1977; Jordan *et al.* 1980; Colombotos 1969; Wiseman 1972), and in other countries such as Denmark (Kormendi *et al.* 1986). Simple straight-forward questions asked identically by telephone and face-to-face yield similar distributions of response.

In the SCPR studies the marginal distributions of response yielded by the different modes of interview were compared and differences were tested for statistical significance using chi-squared tests. These tests were performed on unweighted data. However, tables in the text, unless otherwise indicated, show distributions of data weighted to take account of any differences between the number of people listed on the Electoral Register and those found at an address. Such differences occurred in approximately 25% of cases, in each of which the data were weighted by the number of persons aged 18 or over living at the address divided by the number of electors listed on the Register for that address. Weighted tables are given to allow readers to decide if they might draw different conclusions from telephone survey data and face-to-face survey data when both sets have been prepared according to routine procedures.

Standard chi-squared tests were performed even though the data arose from a multi-stage sample. It has been shown (see, for example, Holt, Scott and Ewings 1980) that underestimating true variability by ignoring sample design will generally lead to test statistics which are too large, and hence to the false rejection of null hypotheses (i.e., to anti-conservative tests). For the Social Attitudes Survey, however, estimation of true standard errors for attitudinal variables yields Design Factors (the ratio of the complex standard error to the simple random sampling standard error) which are rarely above 1.2 (Jowell and Witherspoon 1985). Further, the literature argues that in 2-way tests of independence the consequences of clustering are likely to be less severe (Holt, Scott and Ewings 1980). As a result, we feel justified in using standard chi-squared tests to avoid the large amount of computation necessary for corrected statistics. If anything, this approach will overstate the significance of differences between interview modes.

In the first study we looked at 95 questions and parts of questions and in the second study 69. The results are shown in Table 4. It is clear that in both studies the results accorded with those of other researchers: the interviewing modes yield significantly different distributions of answers for only a very small percentage of questions. A similar finding emerged from the MRDF study.

3.1.2 Comparisons for Particular Question Forms

Despite the general result, research in the U.S. has shown that there are specific kinds of questions for which differences in response distributions do occur. For example, Groves and Kahn (1979) demonstrated a tendency for respondents to give truncated answers to open-ended items over the telephone. This might be due to the faster pace of telephone interviewing, as noted, for example, by Dillman (1970) and Williams (1977). Both interviewers and respondents tend to speak more quickly on the telephone and to avoid silent pauses. The swifter pace of telephone interviews was shown in our second experiment. As Table 5 shows, with an interview designed to take 25 minutes, 10% of the telephone interviews were conducted in under 20 minutes, compared with 5% of face-to-face interviews. At the other extreme, 41% of face-to-face interviews took more than half an hour compared with under a third of the telephone interviews.

Ball (1980) suggests that the greater speed may occur because the norms of telephone conversations require both the interviewer and respondent to work to maintain the conversational flow. This may leave respondents with less time to think about their answers. Certainly, silences seem to make people uncomfortable – in a study by Jordan (1980) routine pauses in the interview were described as interminable by interviewers. Undoubtedly there are many other contributing factors: even the absence of visual distractions may be important.

Although SCPR's experimental studies did not carry any open-ended items, the MRDF study included a number of spontaneous awareness measures. Comparisons of telephone and face-to-face results appear consistent with the findings discussed above. One example

Table 4
Differences in Marginal Distributions of Response:
Telephone vs. Face-to-Face

Bases	Study 1 (95)	Study 2 (69)
	%	%
No significant difference	91	87
Significant at 5%	7	9
Significant at 1%	2	4

Table 5
Interview Length by Mode of Interview (Study 2)

Unweighted Bases	Telephone (354)	Face-to-Face (360)
	%	%
Minutes		
Under 20	10	5
20-29	63	53
30-40	22	33
40+	6	8

$\chi^2 = 17.6$ d.o.f. = 3 $p < 0.01$

Table 6
Comparisons of Responses on an Open Question (MRDF Survey)

Bases	Telephone (700)	Face-to-Face (601)
	%	%
What do you like about . . . soup?		
Number of answers		
None	33	22
One	58	61
Two	7	14
Three or more	1	2
Average	0.77	0.96

$\chi^2 = 32.2$ d.o.f. = 3 $p < 0.01$

is given in Table 6, which shows that a third of telephone respondents gave no answers, compared with under a quarter face-to-face. Also, the average number of responses given over the telephone was significantly lower.

We might assume that more or longer answers mean more valid reporting, and this would imply a need for techniques to improve open questions on telephone surveys. At the extreme, it might be concluded that open questions have only limited use on telephone surveys, for example when only the first information spontaneously offered by respondents is wanted. This assumption needs, however, to be tested: here we can only report the effect.

Differences between response distributions have also been reported for attitude scale questions asked identically face-to-face and over the telephone. Telephone respondents tend towards "acquiescence" and "extremeness" response bias (Jordan, Marcus and Reeder 1980; Groves and Kahn 1979). With the agree/disagree scales used by MRDF, the telephone sample showed a slight tendency to agree more. However, no difference in the spread of responses was found – there was no evidence of a greater tendency towards extremeness.

3.1.3 Sensitive Questions

Concerning the types of question that can be used in telephone surveys, researchers have paid much attention to sensitive questions – those that deal with private or personal information and those for which certain responses are more clearly socially acceptable. Initial views about the likely effects of asking sensitive questions over the telephone were divided. Those who felt that respondents would be less willing to answer truthfully said that the lack of the interviewers' reassuring presence would make respondents less likely to be frank and open. The opposite view – that respondents would give more valid answers – maintained that greater social distance, by preserving anonymity, would encourage truthful responses.

Most evidence supports the latter view (Colombotos 1965; Wiseman 1972; Henson, Roth and Cannell 1974; Locander 1974; Rogers 1976). The major exception is reported by Groves and Kahn (1979), who found telephone respondents to be reticent about their financial status and other sensitive issues.

Our studies support the hypothesis that telephone surveys work well for sensitive questions. For instance, in our first study 14 questions were isolated as potentially sensitive and tested for mode-effects. Three illustrative examples of such questions are given below:

- i) How would you describe yourself?:
(Read out) ...
... as very prejudiced against people of other races
... a little prejudiced
... or, not at all prejudiced?
- ii) Do you think, on the whole, that Britain gives too little or too much help to Asians and West Indians who have settled in this country, or are present arrangements about right?
- iii) Finally in this section, I would like you to tell me whether, in your opinion, it is acceptable for a homosexual person to be a teacher in a school?

No significant differences in the marginal distributions of response were found. For several questions, however, there was a somewhat greater tendency to give socially desirable answers in face-to-face contact. In other words, the questions seemed to be less sensitive over the telephone. For example, 28% of respondents interviewed by telephone admitted to having been questioned by police over the past two years in connection with a crime, compared with 20% of face-to-face respondents.

Sensitive questions in the MRDF study also showed a slight tendency for telephone respondents to give more “honest” answers, although on individual questions differences in the distributions were generally not significant. For example, when asked to describe themselves on a number of dimensions, telephone respondents were more likely to say they were “attractive” (mean score of 2.81 out of 4 compared with 2.72 face-to-face) and were more ready to give an answer at all (88% gave an answer compared with 75% face-to-face).

Questions about income have generally been regarded as potentially problematic in telephone surveys, both in respondents’ willingness to answer and in the answers given. Under-reporting of income levels is the main expectation, although in practice this may be hard to distinguish from under-estimation resulting from higher non-response in the upper income brackets. A study by Locander and Burton (1976) suggests that the validity of income data may depend on the question format. In a comparison of four question formats, under-reporting of income resulted from a method that first asked “Is your income more than \$2,000?” gradually increasing the figure until the first “no” response. However, over-reporting of income was encouraged by a similar method that began with the highest income category. The method used for the telephone surveys in the SCPR experiments was similar to the first type described above. It most closely approximates the response task set by the face-to-face income question in which a card indicating broad income bands, starting with the lowest, was used to guide the respondents’ choice. Over the telephone, the ranges were read to respondents starting at the lowest levels. The results are shown in Table 7.

In neither study was there any mode difference in respondents’ willingness to answer the income question. Differences in the distribution of answers, in this case a possible under-reporting of income, were only apparent in the first study.

3.1.4 Complex Questions

In both SCPR studies a number of questions were identified in advance as likely to pose particular response problems for telephone respondents. These included questions with one or more potentially difficult concepts, long questions and questions with large numbers of response options. Such “complex” questions appear to be no more problematic for telephone respondents than for those interviewed in person. For example, of 19 “complex” questions

Table 7
Gross Household Income: SCPR Studies

Bases ^a Income	Study 1		Study 2	
	Telephone (183)	Face-to-Face (170)	Telephone (297)	Face-to-Face (352)
less than £5,000	38	27	28	28
£5,000–£9,999	42	37	37	38
£10,000 or over	21	35	35	35
Study 1: $\chi^2 = 10.08$ d.o.f. = 2 $p < 0.01$				
Study 2: $\chi^2 = 0.11$ d.o.f. = 2 $p > 0.9$				
Bases	(217)	(199)	(344)	(405)
Don't know/ Not answered	16%	15%	14%	13%

^a “Don't know” and “Not answered” excluded.

identified on the first study (12 of which had been asked with the aid of show-cards face-to-face), only one showed any evidence of mode-effects.

4. SUMMARY AND CONCLUSIONS

Since telephone ownership in the United Kingdom remains relatively low, particularly for certain sectors of the population, telephone interviewing is unlikely to replace face-to-face interviewing for surveys that must include the less advantaged. But its potential in combination with traditional face-to-face procedures has gained recognition. For example, the U.K. Labour Force Survey uses telephone interviewing for second and subsequent interviews with eligible respondents who have indicated a willingness to be contacted by telephone.

Crucial to the success of dual-mode surveys is the absence of differential mode effects. The results reported here provide a largely optimistic outlook. With a few exceptions there were no statistically significant differences between the distributions of answers obtained face-to-face and those given over the telephone.

However, the relatively low response rates to telephone surveys poses problems that need to be overcome. High refusal rates can reduce the cost-effectiveness of using the telephone. More importantly, they increase the chances of introducing bias into the sample. Further research to explore ways of improving telephone response rates is necessary to realize the potential of the method in the United Kingdom.

ACKNOWLEDGEMENTS

This research was carried out within the SCPR Survey Methods Centre, funded as a Designated Research Center by the Economic and Social Research Council (grant HR 3333). The authors are grateful to the editors and referees for their extensive and useful comments on earlier drafts of this paper.

REFERENCES

- ARONSON, S. (1971). The Sociology of the Telephone. *International Journal of Comparative Sociology*, 12, 153-167.
- BALL, D.W. (1968). Towards a Sociology of Telephones and Telephoners. In *Sociology and Everyday Life* (ed. Marcello Truzzi), Englewood Cliffs, New Jersey: Prentice Hall.
- BERGSTEN, J.W. (1979). Some Methodological Results from Four Statewide Telephone Surveys Using Random Digit Dialing. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 239-243.
- BISHOP, Y.M.N., FIENBERG, S.E., and HOLLAND, P.W. (1975). *Discrete Multi-variate Analysis: Theory and Practice*. Cambridge, Mass: MIT Press.
- BLANKENSHIP, A.B. (1977). *Professional Telephone Surveys*. London: McGraw-Hill.
- BLYTH, W.G., and MARCHANT, L.J. (1973). A self-weighting random sampling technique. *Journal of the Market Research Society*, 15, 157-162.
- CANNELL, C.F., OKSENBERG, L., and CONVERSE, J.M. (1979). Experiments in interviewing techniques. Research Report, Institute for Social Research, University of Michigan.
- CHRISTOFFERSEN, M.N. (1984). The quality of data collected at telephone interviews. Danish National Institute of Social Research, Copenhagen.

- COLLINS, M. (1983). Telephone interviewing in consumer surveys. *Market Research Society Newsletter*, October.
- COLLINS, M., and SYKES, W. (1987). The Problems of Non-Coverage and Unlisted Numbers in Telephone Surveys in Britain. *Journal of the Royal Statistical Society*. Ser. A, 150, (forthcoming).
- COLOMBOTOS, J. (1965). The effects of personal vs. telephone interviews on socially acceptable responses. *Public Opinion Quarterly*, 29, 457-458.
- COLOMBOTOS, J. (1969). Personal versus telephone interviews: effect on responses. *Public Health Reports*, 84, 773-782.
- COOMBS, L., and FREEMAN, R. (1986). Use of telephone interviews in a longitudinal fertility study. *Public Opinion Quarterly*, 28, 112-117.
- CZAJA, R., BLAIR, J., and SEBESTIK, J.P. (1982). Respondent selection in a telephone survey: a comparison of three techniques. *Journal of Marketing Research*, 19, 381-385.
- DE MAIO, T.J. (1984). Refusals in telephone surveys: when do they occur? Paper presented at the 39th Annual Conference of the American Association for Public Opinion Research.
- DILLMAN, D. (1970). *Mail and Telephone Surveys: The Total Design Method*. New York: John Wiley.
- DILLMAN, D., GALLEGOS, J., and FREY, J. (1976). Reducing refusal rates for telephone interviews. *Public Opinion Quarterly*, 40, 66-78.
- FALTHZIK, A. (1972). When to make telephone interviews. *Journal of Marketing Research*, 9, 451-452.
- FITTI, J.E. (1979). Some results from the telephone health interview survey. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 244-249.
- FLEISHMAN, E., and BERK, M. (1979). Survey of interviewer attitudes towards methodological issues in the national medical care expenditure survey. Paper presented at the Third Biennial Conference on Health Survey Research and Methods, Reston, Virginia.
- GROVES, R., and KAHN, R. (1979). *Surveys by Telephone: A National Comparison with Personal Interviews*. New York: Academic Press.
- GROVES, R.M., MAGILAVY, L.J., and MATHIOWETZ, N.A. (1981). The process of interviewer variability. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 438-443.
- GROVES, R.M., and MATHIOWETZ, N.A. (1984). Computer assisted telephone interviewing: effects on interviewers and respondents. Survey Research Center, University of Michigan.
- HENSON, R., ROTH, A., and CANNELL, C.F. (1974). Personal vs. telephone interviews and the effects of telephone re-interviews on reporting of psychiatric symptomatology. Research Report, Survey Research Center, University of Michigan.
- HOCHSTIM, J.R. (1967). A critical comparison of three strategies of collecting data from households. *Journal of the American Statistical Association*, 62, 976-986.
- HOLT, D., SCOTT, A.J., and EWINGS, P.D. (1980). Chi-squared tests with survey data. *Journal of the Royal Statistical Society*, Ser. A, 143, 303-320.
- IBSEN, C.A., and BALLWEG, J. (1974). Telephone interviews in social research: some methodological considerations. *Quality and Quantity*, 7, 181-192
- JOWELL, R., and WITHERSPOON, S. (1985). *British Social Attitudes: The 1985 Report*. Aldershot: Gower.
- JORDAN, L.A., MARCUS, A.C., and REEDER, L.G. (1980). Response styles in telephone and household interviewing: a field experiment. *Public Opinion Quarterly*, 44, 210-222.
- KAHN, R.L., and GROVES, R.M. (1977). *Comparing telephone and personal interview systems*. Survey Research Center, University of Michigan.
- KORMENDI, E., EGSMOSE, L., and NOORDHOEK, J. (1986). Datakvalitet ved Telefon-interview. Socialforskningsinstituttet, Studie 52, Copenhagen.

- LOCANDER, W.B., and BURTON, J.P. (1976). The effect of question form on gathering income data by telephone. *Journal of Marketing Research*, 13, 189-192.
- LOCANDER, W.B., SUDMAN, S., and BRADBURN, N. (1974). An investigation of interview method, threat and response distortion. *Proceedings of the Social Statistics Section, American Statistical Association*, 21-27.
- LUCAS, W.A., and ADAMS, W.C. (1977). *An Assessment of Telephone Survey Methods*. Santa Monica, California: Rand Corporation.
- McCULLAGH, P., and NELDER, J.A. (1983). *Generalised Linear Models*. London: Chapman and Hall.
- MILLER, P.V., and CANNELL, C.F. (1982). A study of experimental techniques for telephone interviewing. *Public Opinion Quarterly*, 46, 250-269.
- Market Research Development Fund, (1985). *Comparing telephone and face-to-face surveys*. Marplan Ltd.
- OKSENBERG, L., COLEMAN, L., and CANNELL, C. (1984). Voices and refusal rates in telephone surveys. Unpublished manuscript.
- O'NEIL, M., GROVES, R., and CANNELL, C. (1979). Telephone interview introductions and refusal rates: experiments in increasing respondent cooperation. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 252-255.
- ROGERS, T.F. (1976). Interviews by telephone and in person. *Public Opinion Quarterly*, 40, 51-65.
- SCHMIEDESKAMP, J.W. (1962). Reinterviews by telephone. *Journal of Marketing*, 26, 28-34.
- SYKES, W., and HOINVILLE, G. (1985). Telephone interviewing on a survey of social attitudes: a comparison with face-to-face procedures. Social and Community Planning Research Center.
- TODD, J., and BUTCHER, R. (1982). *Electoral Registration in 1981*. London: OPCS.
- WILLIAMS, E. (1977). Experimental comparisons of face-to-face and mediated communication. *Psychological Bulletin*, 84, 963-976.
- WISEMAN, F. (1972). Methodological bias in public opinion surveys. *Public Opinion Quarterly*, 34, 105-108.
- WISEMAN, F., and McDONALD, P. (1979). Noncontact and refusal rates in consumer telephone surveys. *Journal of Marketing Research*, 16, 478-484.