# The Use of CAPI for Attitude Surveys: An Experimental Comparison with Traditional Methods

Jean Martin[1], Colm O'Muircheartaigh[2], and John Curtice[3]

**Abstract:** Three split sample comparisons of interviews using laptop computers and traditional pencil and paper methods were carried out on surveys consisting largely of attitude questions. Two of the experiments were conducted on waves two and three of a panel survey; the third used a fresh sample of respondents. CAPI was found to be acceptable to both interviewers and respondents; no differences in either initial response rates or response following a computer interview were found. The increased length of CAPI interviews on two of the studies was attributed to interviewers' lack of experience.

Comparison of responses to attitude questions showed no difference in use of "don't know" or midpoint options, but CAPI respondents were more likely to choose the extreme responses of five and eleven point scales. There was some evidence that CAPI might improve stability of responses over time and none that the technique would reduce it.

**Key words:** Computer assisted interviewing; data quality; response rates; response effects.

## 1. Background

### 1.1. Computer assisted telephone and face-to-face interviewing

Computer assisted interviewing (CAI) has

been in use now for well over a decade. Most of the early developments, particularly in the USA, were undertaken on telephone surveys – computer assisted telephone interviewing (CATI). There are though, in Great Britain at least, serious impediments to the use of the telephone on surveys which are designed to be representative of the general population. Not only do 12% of households in Great Britain currently not have telephones but also the members of these households differ in many important respects from telephone owners. The elderly, the unemployed, single parents and other low income groups are all under-represented amongst telephone owners. In addition there are still major unresolved methodological obstacles to the successful selection and contact of an

adequate probability sample of telephone owners while response rates also tend to be lower than on face-to-face surveys.

Thus in Great Britain surveys which require fully representative samples will continue to require a component of face-to-face interviewing for the foreseeable future. But the possibility of incorporating some of the advantages of computer assisted methods into the face-to-face interviewing has been opened up by the advent of light and cheap portable computers capable of supporting a computer assisted interview carried out in respondents' homes.

Computer assisted personal interviewing (CAPI) is still in its infancy. This is particularly true in countries such as the USA where CATI is widely used and where in consequence CAPI is regarded as a recent offshoot from CATI. But the use of CAPI is still limited even in countries such as The Netherlands, Sweden and Great Britain where CATI has not been traditionally used on surveys requiring representative samples and where CAPI is seen as an alternative to traditional paper-and-pencil (PAPI) techniques rather than as an extension of CATI. In Britain CAPI is currently used on only two major government surveys, the British Labour Force Survey carried out by the Office of Population Censuses and Surveys (OPCS), and the new Family Resources Survey, carried out jointly by OPCS and Social and Community Planning Research. So far just one commercial organisation is equipped to carry out major CAPI surveys but others are likely to follow soon.

## 1.2. The nature and advantages of computer assisted interviewing

The questionnaire for a CAI survey is specified in a computer program which tells the system about the form of the questions and answers, the range of permissible responses, and the routing instructions which determine which respondents should be asked which questions. Checks on the consistency of responses, to be carried out while the interview is in progress, can also be specified. During the interview the questions and any other instructions are displayed on the computer screen. The interviewer reads out the current question and keys in the appropriate response. The computer checks the validity of the response and, if and when an acceptable response has been registered, then selects and displays the next question. Text answers may still need to be coded after the interview, but other processes traditionally needed with PAPI, such as specifying the variables for the analysis package and keying and editing the data, are eliminated or greatly reduced.

The potential advantages of CAI are considerable:

– *Reduction in the time elapsed between the fieldwork and the availability of the data for analysis*. At the end of the interview the data are in electronic form and, apart from the coding of text answers, are ready to be combined into a simple raw data set for analysis. Some CAI programs will produce ready-labelled system files for SPSS, SAS, etc. Thus for surveys where timeliness of results is important, CAI has much to offer.

– *Improvement in data quality*. Since interviewers are routed by the computer to the next appropriate question, routing errors by the interviewer are impossible. Answers outside the specified range cannot be entered. Consistency checks can be programmed to be carried out during the interview rather than later in the office; this has the advantage of allowing the interviewer to resolve inconsistencies directly with the respondent.

– *Reduction in survey running costs*. CAPI avoids the cost of producing and handling paper questionnaires as well as most if not all of the data keying and office editing normally required after interviewing. These savings in current costs do however have to be balanced against higher *initial hardware, software and other set-up costs*. These mean that it is likely to be some time before net savings in survey costs become apparent. But on large continuing surveys where the set-up costs can be amortised over a number of years, CAI is economically attractive.

## 1.3. Evidence from other studies

While CAPI may be economically most suited for continuing long-term surveys, it is also precisely on such surveys that there is the greatest need to be sure that a change in mode of administration will not affect the pattern of response. There is an extensive literature on the modal effects associated with face-to-face, postal and telephone questionnaires; on the modal effects associated with telephone interviewing in particular, see Groves and Mathiowetz 1984; Groves and Nicholls 1986; Catlin and Ingram 1988. Yet, in contrast, there has been to date only limited study of CAPI. Furthermore, most of the literature is largely descriptive in character and there have been few properly controlled studies comparing computer assisted with paper and pencil interviewing.

What work has been conducted has been largely encouraging. Four propositions appear to be supported by the existing literature:

1. Interviewers are able to carry out CAPI interviews successfully once they have been suitably trained. Interviewers find CAPI at least as acceptable as PAPI and indeed in some table as PAPI and indeed in some

studies have indicated a preference for CAPI (van Bastelaer, Kerssemakers, and Sikkel 1988; Manners 1987; Sikkel 1988; Baker 1990; Blom 1990; Bradburn, Frankel, Baker, and Pergamit 1991; Edwards, Edwards, Gay, and Sperry 1992). Survey organisations have been able to train their ordinary face-to-face interviewers and have not needed to recruit different types of interviewer for CAPI.

2. Response rates are no lower to CAPI than to PAPI administered interviews (van Bastelaer et al. 1987; Manners 1987; Sikkel 1988; Blom 1990; Bradburn et al. 1991). When asked their reactions to CAPI, respondents range from "indifference to enthusiasm" (Baker 1990) while no negative reactions were reported. Indeed, when the National Opinion and Research Center (NORC) asked respondents directly about their preference for CAPI or PAPI, most said they preferred CAPI.

3. Computer assisted interviewing improves data quality. Both the CATI and the CAPI literature report fewer instances of missing data thanks to the impossibility of routing errors (Groves and Mathiowetz 1984; Groves and Nicholls 1986; Catlin and Ingram 1988; Sebestik, Zelon, De Witt, O'Reilly, and McGowan 1988). But in addition the incidence of "don't know" responses and refusals has also been reported to be lower on CAPI surveys (Baker and Bradburn 1991; Bradburn et al. 1991).

4. Although CAPI is not suitable for surveys which contain large numbers of open questions, the quality of the recorded text is no lower on surveys which contain one or two open questions (such as descriptions of

occupation and industry) (INSEE experiment summarised in Baker 1990; Manners, personal communication, 1990). The fact that few interviewers are competent typists does not appear to be an impediment.

5. CAPI may increase respondents' confidence in the confidentiality of their answers and thus make them willing to give fuller and perhaps more accurate reports of socially unacceptable behaviour (Baker and Bradburn 1991; Bradburn et al. 1991).

But although encouraging, the evidence is not yet sufficient to enable one to introduce CAPI into a *non-governmental* time series in which *attitudinal* questions predominate. From the perspective of such a series, there are three substantial reasons for caution:

1. Little of the research has been undertaken on surveys which contain a high proportion of attitude questions. Yet it is known that attitude questions are generally more sensitive to changes in administration or methodology than are factual questions.

2. Many of the surveys on which CAPI work has been undertaken have been government surveys, such as the labour force surveys in the UK, Netherlands, USA, Sweden and Canada. The level of response to such surveys is generally higher than on non-governmental surveys and thus any resistance to the introduction of CAPI may be less likely to manifest itself than on non-governmental surveys.

3. There is some evidence that it takes longer to administer a CAPI interview than a PAPI interview (Couper, Groves, and Kosary 1989), although this finding was not replicated on a

recent NORC study (Bradburn et al. 1991).

## 2. Aims of the Study

This paper is intended to address some of these remaining gaps and concerns. It reports the findings of three experiments into the possible effect of CAPI conducted on two non-governmental survey projects both of which contained a large number of attitude questions. In the first half of the paper we consider how far our evidence confirms or otherwise the applicability of existing findings on CAPI to non-governmental surveys. We consider four topics:

i. the acceptability of CAPI to interviewers,

ii. the acceptability of CAPI to respondents,

iii. the effect of CAPI on the length of interview,

iv. the effect of CAPI on the levels of missing data.

In the second half of the paper we report our findings on the possible effect of CAPI on responses to attitude questions. In particular we consider:

v. the patterns of response to attitude questions, and

vi. the stability of responses over time.

In so doing our main aim is to consider whether or not the introduction of CAPI into attitude surveys might compromise the integrity of a time series.

## 3. Design of the Experiments

In each of the three experiments the relevant samples were split into two halves; half of the respondents were interviewed by CAPI and half by PAPI. Two of the experiments were undertaken as part of a programme of research of the Joint Unit for the Study

of Social Trends (JUSST). The experiments (referred to hereafter as JUSST1 and JUSST2) took place on waves two and three of a three wave panel survey of respondents who were first interviewed (with PAPI) as part of the 1989 British Social Attitudes (BSA) survey. (The BSA series of surveys is an annual series of cross-section surveys inaugurated in 1983 by Social and Community Planning Research, London. The survey, which is modelled on NORC's General Social Survey in the United States, is designed to track trends in social, political and economic attitudes amongst the general population of Great Britain.) The main purpose of the JUSST survey was to improve and develop questions to measure social and political values. Many of the questions included have been used or are intended for use on either the BSA surveys or the British Election Studies. (The British Election Study series is a series of cross-section and panel surveys conducted at each British general election since 1964, and is the longest running academic social survey in Great Britain. It is designed to study influences on voting behaviour and the formation of social attitudes, and performs a similar role to that of the American National Election Study at the University of Michigan in the United States.) The fact that the experiments took place as parts of a larger project placed some constraints on the way the experiment was conducted.

The third experiment – the Election Study Methodology (ESM) project – was undertaken wholly for the purpose of comparing results from CAPI and PAPI interviews. It was financed by the Economic and Social Research Council in order to inform the design of future British Election Studies.

All three experiments have their limitations. They covered just 397 and 348 respondents on JUSST1 and JUSST2,

respectively. The overall response rates on the JUSST1 and JUSST2 surveys were 73% and 80%, respectively, meaning that just 58% of those originally selected for inclusion were interviewed at JUSST2. Respondents who had indicated their unwillingness to be re-interviewed after the original 1989 BSA interview were not approached and for practical reasons no interviewing was conducted in Greater London. The ESM survey covered as many as 640 respondents from across the whole of Great Britain but this represented a response rate of just 52%.

The design of the experiments was also constrained by costs; it was possible to purchase only 10 laptop computers for the project. In order to maximise the number of interviewers taking part 10 interviewers used the computers for the first half of the fieldwork period while a different 10 interviewers used them for the second half. Thus 20 interviewers took part in each experiment. This design controls for any effect due to which mode was used first but has the disadvantage that interviewers had less time to complete their CAPI interviews than their PAPI interviews. Consequently a few interviews which were assigned to the CAPI sample and where the interviewer had access to a computer during the first half of the fieldwork were in fact carried out by PAPI because the interviewer had been unable to make contact in the first fieldwork period.

But our experiments (of which full details are given in the appendix) also have two methodological strengths. Firstly, for each experiment each interviewer covered one or more sampling points and within each sampling point respondents were allocated randomly to CAPI or PAPI. Each interviewer was allocated equal numbers of cases for each mode. This means that any difference between the two modes cannot

be a consequence of interviewer differences. In addition, this strategy meant that because both samples came from the same sampling points, the likelihood that there would be substantial differences in the socio-economic characteristics of CAPI and PAPI respondents was reduced.

Secondly, different respondents were selected for the JUSTT1 and JUSST2 experiments. Furthermore, with just two exceptions, different interviewers took part in these two experiments. This means that we can examine the possible effect of CAPI on the stability of responses over time. This would reveal itself in a difference in the stability of the responses of those who were interviewed by CAPI at either JUSST1 or JUSST2 compared with those who were only ever interviewed by PAPI.

## 4. Reactions of Interviewers

For all the interviewers working on JUSST1, and all but two on JUSST2, this was their first experience of CAPI. After carrying out their interviews, interviewers were asked to complete a short questionnaire about their reactions – and their perceptions of their respondents' reactions – to the experience of interviewing by CAPI. Most of the reactions were positive. Sixteen out of 19 interviewers on JUSST1 said that, overall, they enjoyed using the laptop and

would like to use it again. On JUSST2 interviewers were more evenly divided between those preferring CAPI, those preferring PAPI and those with no preference.

On both JUSST experiments several interviewers mentioned problems of maintaining rapport with respondents because they were concentrating on the technical aspects of using the computer. But such problems were reported less frequently on the ESM study where all but three of the interviewers had worked on one of the JUSST experiments, suggesting they were a function of lack of experience. Similarly, although few problems affected the actual interviewing, several interviewers contacted HQ with technical difficulties and queries during the JUSTT surveys. But again far fewer problems were reported on the ESM study again suggesting they were teething problems. Our study therefore confirms existing findings that there is every reason to believe that interviewers are able to handle CAPI, and that indeed in many cases they find it sufficiently attractive to want to use computers again.

## 5. Reaction of Respondents

What about respondents themselves? Did they react adversely to the use of a computer for interviewing? Despite their widespread

*Table 1.   Indicators of respondent reactions*

|  | JUSST1 | | JUSST2 | | ESM | |
|---|---|---|---|---|---|---|
|  | CAPI | PAPI | CAPI | PAPI | CAPI | PAPI |
| Refusal rate | 14% | 13% | 9% | 11% | 23% | 22% |
| Definitely or maybe willing to be reinterviewed | 94% | 91% | – | – | 89% | 87% |
| No preference for CAPI vs. PAPI | 82% | – | 82% | – | 80% | – |
| No. of respondents | (138) | (152) | (131) | (147) | (163) | (170) |

use, many members of the public will not have had much contact with computers and it is possible that a general fear or mistrust of them might affect their willingness to be interviewed by this mode, particularly on a non-governmental survey. We therefore examine respondents' responses to being approached for a CAPI compared with a traditional paper interview and also whether the experience of a CAPI interview affects willingness to participate in a further interview.

## 5.1. Response to approach to interview

Non-response can arise for a number of reasons – only some of which can possibly have anything to do with the mode of interview. Clearly if no contact has been made with a respondent because he or she was persistently out or had moved, the use of a computer can scarcely have affected the outcome. What may be influenced is the level of refusals. This level (including those who broke appointments) is shown for each experiment in Table 1.

The ESM survey was undertaken on a fresh sample while respondents to the two JUSST surveys had already been interviewed once before and had agreed to be approached again. So, not surprisingly, the overall refusal rates to the two JUSST surveys are lower than on the ESM survey. But there is virtually no difference in the response rate between the two modes on each survey. Our results confirm that the introduction of CAPI has no overall effect on response rates – and that this is true of panel as well as cross-section surveys.

Despite the lack of overall difference, different types of respondents might still have reacted differently to the two modes of interview. Younger people might be more familiar with new technology and less scared of computers than older people,

and thus more likely to agree to a computer interview; on the other hand greater familiarity with computers might make younger people more suspicious of how computer data might be used and thus less likely to agree to a computer interview. Equally, it is sometimes argued that men are more interested than women in new technology. For the JUSST surveys we had information on the demographic characteristics of respondents and non-respondents from their 1989 BSA interview, so we were able to investigate such possibilities. Analysis showed that there were no differences in response according to mode of interview by either gender or age.

## 5.2. Willingness to be re-interviewed

Respondents may, of course, develop an adverse reaction to being interviewed by computer after they have experienced it. One way of identifying this is to examine the answers given to a request for permission to conduct a subsequent interview. Such a question (which is commonly asked in cross-section as well as panel surveys) was asked at the end of both the JUSST1 and ESM interviews. If respondents did not like being interviewed by computer or found the interview more of a burden than a paper interview, we would expect them to be less willing to agree to a further interview. But Table 1 shows there was no significant difference in willingness to be re-interviewed between the two modes. Equally, further analysis revealed no differences according to either age or gender.

Agreeing to a further interview is one thing; the acid test is whether respondents actually respond to a further approach. The panel design of the JUSST experiments allowed us to test just this. Of those who had been interviewed by CAPI at JUSST1, 77% responded at JUSST2, the corresponding

response rate for those interviewed by PAPI being 76%. In short, there seems no reason for concern that the introduction of CAPI would have an adverse effect on response to further waves of a panel survey.

### 5.3. Views of respondents

Respondents were asked their feelings about being interviewed using a computer at the beginning and the end of each CAPI interview. At the beginning most respondents said they felt interested but a few were a little suspicious and apprehensive about computers in general. Some thought that the use of computers made the interviewing process impersonal but that it would probably save time and, because no names were entered on the computer file, guard confidentiality. When asked at the end of the interview whether they would prefer to be interviewed using a computer or an ordinary paper-and-pencil questionnaire, the majority of respondents expressed no preference (see Table 1). Of those who did, more favoured CAPI than PAPI.

### 5.4. Conclusions

These results confirm those of other studies which show that CAPI is just as acceptable as PAPI for most respondents. The introduction of computers does not reduce willingness to respond positively to a first request for an interview and neither does the experience of CAPI affect propensity to take part in subsequent waves of a panel study.

### 6. Length of Interview

We remarked earlier that the existing literature has produced contradictory findings on whether CAPI increases interview length. There are clear reasons why it might do so. Most notably, interviewers cannot read ahead; they have to enter the response for the current question before they can read the next question. And in recording the answers to open questions we have to remember that even if they have some keyboard skills, few interviewers can type as fast as they can write. On the other hand any differences we might find could be a function of experience rather than a modal difference. All of our interviewers were experienced in the use of PAPI; none were so in the use of CAPI. Thus one might anticipate they would take longer to complete CAPI interviews initially while they were learning new skills. In particular, proficient use of a computer requires keyboard skills which interviewers may not have at all; it was noticeable at the training classes that interviewers who had never used a typewriter before were much slower than other interviewers as they had to find the standard keyboard keys as well as the special computer keys.

As Table 2 shows, CAPI interviews did indeed take longer on both JUSST surveys ($p < .001$) when all of the interviewers were using CAPI for the first time. On average the difference was about seven minutes, though it varied considerably from one interviewer to another. In contrast, on

Table 2.    *Average length of interview (minutes) by mode*[1]

|  | JUSST1 | | JUSST2 | | ESM | |
|---|---|---|---|---|---|---|
|  | CAPI<br>($N = 135$) | PAPI<br>($N = 151$) | CAPI<br>($N = 130$) | PAPI<br>($N = 147$) | CAPI<br>($N = 158$) | PAPI<br>($N = 165$) |
| Mean (mins) | 60.9 | 54.1 | 55.9 | 48.4 | 47.6 | 45.4 |
| S.D | 13.2 | 12.8 | 14.6 | 12.2 | 15.2 | 14.2 |

[1]Excludes partial and implausibly long interviews.

the ESM study – where more of the interviewers had some experience of CAPI – the CAPI interviews took on average just two minutes longer, a difference that was not statistically significant.

The results of the JUSST experiments need, however, to be considered in a little more detail. They could be the result of a set of technical artifacts. For example, there was clear evidence that in recording the interview length most PAPI interviewers rounded to the nearest five minutes. Times for the CAPI interviews in contrast were derived from the computer's internal clock.

Not that these computer timings were unproblematic. Their accuracy relies on the interviewer exiting from the programme as soon as the interview is concluded. There is a simple reason why this may not have always happened. At the end of each CAPI interview, interviewers asked a few additional questions about the respondent's reactions to the use of a computer. They should have turned the computer off before doing this but we cannot guarantee that they always did so. If they did not, the time taken to administer these extra questions would have been included in the recorded interview length.

In addition, in a few cases the length of the interview calculated from the times given for the start and end of the interview were implausibly long – perhaps because the interview had been interrupted. These cases have been excluded from Table 2.

The lessons of this experience were taken on board in conducting the ESM study. For example, we recorded the time at various points in the interview rather than just at the beginning and end. And as we have seen the overall CAPI and PAPI interview lengths were not significantly different.

There were though some differences in the time taken for different parts of the ESM questionnaire. Altogether, four sections of the questionnaire were separately timed. In addition, we were concerned at how long three complex scale questions would take and also recorded the time at the start and finish of each of these. These questions were all asked during the third section of the questionnaire but were separated by other questions.

There were in fact no differences between modes in the time taken to complete the first three sections (see Table 3). The differences between the two modes were confined to the remaining parts of the questionnaire.

First of all, CAPI interviews took less time (in two cases significantly less) to administer the three complex scale questions. These had in fact been questions which we thought might be difficult to adapt successfully for computer interviewing. This clearly was not the case – although the reason why they actually took less time probably had more to do with the adoption of a slightly different way of administering the questions in the computer interviews, than because of any inherent modal difference.

The questions required respondents to fill in a booklet which asked them to place themselves and the three main British political parties on eleven-point scales with respect to three ideological dimensions. In the paper interviews, interviewers read out each question and the respondent ticked the box in the booklet to indicate his or her answer. At the end of each set of four questions respondents were asked to review all their answers together, to make any changes they wanted, and then read out the four answer codes which the interviewer recorded in the questionnaire. In the computer interviews, in contrast, respondents told the interviewer the answer code for each individual question as they. went along and the interviewer entered this into

*Table 3.    Time taken for different sections and scale questions: ESM study. The main cell entry is the average time in minutes a section took to complete. The entry in brackets is the standard deviation*

|  | CAPI (Min. $N = 158$) | PAPI (Min. $N = 164$)[1] |
|---|---|---|
| Section A | 5.7 (2.3) | 5.8 (2.8) |
| Section B | 4.3 (1.9) | 4.2 (2.0) |
| Section C | 24.2 (8.2) | 23.6 (8.2) |
| Section D | 13.1 (4.4) | 11.9[2] (4.4) |
| Scale question 1 | 2.6 (1.4) | 2.8 (2.2) |
| Scale question 2 | 1.7 (1.1) | 2.2[2] (1.4) |
| Scale question 3 | 1.4 (0.8) | 1.9[2] (1.2) |

[1] These are the minimum $N$ upon which any cell entry is based. The actual $N$s vary due to item non-response.
[2] Difference between means significant $p < .05$.

the computer straight away. Respondents could still review their answers at the end and make any changes they wanted, but in general this alternative procedure appears to have quickened the interview process. We seem to have here an example of the kind of unintended modal difference brought about by a change in the arrangement of a question when administered by CAPI previously noted by Blom (1990) and Bradburn et al. (1991). There was also a noticeable practice effect on these questions, with the last two booklets administered by each interviewer taking less time than the two previous booklets. The variance of the time taken also decreased with each booklet. The reduction in time due to practice was also more marked for the CAPI than the PAPI sample, and this contributed to the significant differences by mode for the two later scale questions.

Meanwhile, section D took longer when administered by computer. This was the section which included all the questions on background characteristics of the respondents. These included details of the occupation of respondent and spouse – which had to be entered as free text. Here keyboard inexperience seems to have been crucial.

Thus on the basis of our findings there seems no reason to anticipate that, once interviewers have had a small amount of experience, computer interviews will take significantly longer or shorter than paper and pencil interviews – so long as interviewers do not have to enter too much free text.

## 7.    Responses to the Questions

Although our principal and distinctive interest is in attitudinal questions, two of

the experiments contained a number of factual questions for the most eliciting information about respondents' socio-economic background. In addition two of the surveys also included a number of questions designed to test people's knowledge of politics rather than their opinions. So before turning to our attitudinal questions we look at whether the use of CAPI improved the quality of these factual and knowledge questions.

### 7.1. Errors and non-response for factual questions

As explained earlier (see Section 1.2) one of the advantages of CAPI is that interviewers do not have to follow routing instructions. As a consequence interviewers cannot commit routing errors or otherwise omit questions by mistake. The extent to which such errors occur on PAPI surveys depends very much on the complexity of the survey. Surveys such as the BSA and Election Surveys, on which the questionnaires for these experiments were modelled, are not very complex and so few such errors would be expected. Equally, other kinds of item non-response such as a "don't know" response or refusal to answer a question are usually quite infrequent for factual questions of the sort included on such surveys as BSA and the Election Surveys.

Nevertheless our study does provide some limited support for claims that CAPI produces better quality factual data in terms of completeness than PAPI. Although identical on JUSST1, both refusals and don't knows were more numerous in PAPI interviews on the ESM study than in CAPI interviews. Meanwhile, as anticipated, no questions were inadvertently omitted in any of the CAPI interviews.

### 7.2. Don't know responses to knowledge questions

The knowledge questions, which were included on both of the JUSTT surveys, consisted of a short quiz in which respondents were asked whether they thought each of a series of statements was true or false, or whether they did not know. There were 14 statements in the quiz at JUSST1 and 15 at JUSST2.

The crucial question here is whether the use of a computer encouraged respondents to say that they did know the answer and discouraged them from guessing the answer – or the opposite. Overall it did seem that respondents were using the "don't know" option rather than guessing since the proportion of "don't know" answers was negatively correlated with the proportion of correct responses. And mode made no apparent difference. The mean number of don't know responses per respondent was identical for the two modes on both surveys: 2.9 on JUSST1 and 2.2 on JUSST2. The use of a computer apparently does not

*Table 4. Missing data for factual questions by mode of interview*

|  | JUSST1 | | ESM | |
|---|---|---|---|---|
|  | CAPI ($N = 138$) | PAPI ($N = 152$) | CAPI ($N = 160$) | PAPI ($N = 167$) |
| Number of instances of: |  |  |  |  |
|   No answer | 0 | 6 | 0 | 8 |
|   Refused | 5 | 4 | 14 | 25 |
|   Don't know | 0 | 0 | 17 | 25 |
| No. of questions | 22 | | 20 | |

affect propensity to admit to not knowing the correct answer to a factual question in either direction.

### 7.3.  *Responses to attitude questions*

Now we turn to the main concern of this paper – whether or not the introduction of computer assisted interviewing affects the distribution of responses to attitude questions. First of all in this section we compare the distribution of responses for the CAPI and PAPI samples separately for each experiment. Then in the following section we use the panel design of the JUSST surveys to look at stability of responses over time.

The majority of attitude questions were presented in the form of composite questions with a main question stem which was asked of a number of different items. For example, the main question stem "Please choose a phrase from this card to say how you feel about ..." was asked in turn about the Conservative, Labour, Social and Liberal Democrat and Green Parties. The phrases on the card were: "strongly in favour," "in favour," "neither in favour nor against," "against," "strongly against." Because of the relatively small sample sizes involved, and because of the need to carry out multiple comparisons in looking for differences between modes, we have generally analysed each group of sub-questions within a composite question together rather than each question separately. This increases the power of the statistical tests because the stability introduced by grouping similar questions reduces the variance of the comparisons.

Item refusal and omission in error on the attitude questions is too uncommon to permit any useful analysis of modal differences in their frequency. So our attention focuses first of all on the incidence of don't knows, then on the use of the midpoint response

and finally on the incidence of extreme responses.

#### 7.3.1. "Don't know" and neutral responses

As has been the practice on the BSA surveys and recent Election Studies, a "don't know" response code was provided for each question which the interviewer could see but whose presence was not apparent to the respondent. On both the paper questionnaire and the computer screen this response was shown in brackets indicating that the interviewer should not read it out.

Such responses were not common – typically about 1% to 2% of responses per question and only exceptionally more than 5%. To construct a stable measure of their incidence we computed the mean number of "don't know" responses per respondent for each group of sub-questions. This produced 42 separate comparisons, none of which produced a significant modal difference.

Thirty-four of these groups consisted of questions with an odd number of response options (as in 11, 5 and 3 point scales). Here respondents might use the middle category to indicate no opinion rather than give a "don't know" response. It can be seen as another way of giving a neutral response rather than committing oneself to a definite opinion. But again our analysis failed to find any real modal effect. Only one of the groups showed a significant difference between modes in the *combined* occurrence of midpoint and "don't know" responses. Given that when carrying out as many significance tests as this, one in twenty results will be significant by chance anyway (assuming a 5% confidence interval is used). Thus it is unlikely that this one case indicates a real modal effect.

#### 7.3.2.  Mean response

Since we found no difference in the likelihood of respondents choosing a neutral

response, any differences in mean response to the scale questions according to mode of interview indicates a bias in response. Here exceptionally it is appropriate to consider individual questions rather than groups, in order that any biases do not cancel one another. This meant carrying out *t*-tests between the means of some 250 pairs of questions. Only ten of the differences were significant, well within the number that would be expected by chance using a 5% confidence interval. There is therefore no evidence that CAPI results in a bias in responses to scale questions relative to PAPI.

### 7.3.3. Extreme responses
Even though there are no differences in the mean response, the mode of an interview might affect respondents' propensity to choose the extreme responses to a scale question. So we compared the incidence of codes 1 and 5 in the case of 5-point scales, and of codes 1, 2, 10 and 11 in the case of 11-point scales, for the two modes. All three studies give some indication that CAPI respondents are more likely than PAPI respondents to choose an extreme response category. On the JUSST1 survey extreme responses are more common amongst CAPI respondents for six of the

seven attitude scales, although none of the differences are statistically significant. In both the JUSST2 and ESM studies, extreme responses were more common amongst CAPI respondents on all seven and nine attitude scales, respectively. On the ESM survey one of the differences is almost significant at the 5% level. But on JUSST2 three of them were significant, as was the difference in the incidence of extreme responses across the survey as a whole (see Table 5).

Taken individually the results of each of the studies are no more than suggestive of a modal effect. But when all three studies are considered together the evidence is striking. Extreme responses were more common amongst CAPI respondents on no less than 21 of the 22 scales. A non-parametric test of the differences would clearly be significant.

These results might be an artefact of a difference in the character of the respondents included in the two samples. But while there were some differences in the characteristics of CAPI and PAPI respondents in each of the three surveys, the direction of the differences were not consistent across all of the surveys – unlike the differences found in the incidence of extreme responses. We

*Table 5. Average number of extreme responses (codes 1 or 5) to groups of questions with five response categories at JUSST2*

|  | No. of questions | CAPI ($N = 130$) | PAPI ($N = 147$) | Difference |
|---|---|---|---|---|
| 5 point scales: |  |  |  |  |
| Environmental protection | 5 | 1.08 | .74 | .34* |
| Feelings about Britain | 10 | 1.48 | .95 | .53* |
| Feelings about parties | 4 | .89 | .62 | .27* |
| Eastern Europe | 4 | .40 | .29 | .11 |
| Women's roles | 9 | .87 | .63 | .24 |
| Women MPs | 11 | .78 | .65 | .13 |
| Things in common with other groups | 6 | .97 | .78 | .19 |
| All 5 point scales | 49 | 6.47 | 4.66 | 1.81* |

*significant $p < .05$

therefore cannot discount the possibility of a real modal effect.

### 7.4. Conclusions

CAPI has clear advantages in terms of data quality in that routing errors and questions missed accidentally do not occur; the incidence of refusals and "don't know" responses to factual questions also appears to be lower. At the same time, with but one exception, CAPI does not appear to influence the pattern of response to knowledge or attitudinal questions. However, we have found some tendency for respondents to be more likely to give extreme responses to attitudinal questions when interviewed by CAPI.

### 8. Reliability of Responses over Time

Even if there were no differences at all between the two modes of interviewing so far as the distribution of responses is concerned, it could still be possible that the reliability of the responses over time might differ.

Included in the JUSST surveys were a number of questions which were designed to measure respondents' fundamental beliefs and values. The expectation is that these beliefs and values would not vary very much over time – and certainly not in the six-month interval that elapsed between each wave of interviewing. Thus any observed instability will be mainly a reflection of measurement unreliability (error variance). One such source of instability could be a switch from one mode to another between waves of interviewing.

Our basic methodology is to compare the stability of responses of those who were interviewed on both occasions by paper and those who were interviewed on one occasion by paper and on the other by

computer; however, before we can do so we have to bear a number of points in mind. The stability of the responses in the two groups might differ because of differences between interviewers. But the design of the experiments means that we can compare cases in the two sets of laptop areas where both CAPI and PAPI interviews were conducted by the same groups of interviewers. Similarly the aggregate volatility of responses might change between adjacent waves because of differences in the attitudinal stability of those re-interviewed and those not re-interviewed in each successive panel wave. To exclude this our analysis is restricted to those respondents who were interviewed in all three waves (BSA, JUSST1 and JUSST2) of the JUSST panel.

There are still other sources of instability which we cannot control. There could be interaction effects due to changes in the pairing of interviewers and respondents. It is too difficult to exclude such effects by having a particular respondent interviewed by the same interviewer on all occasions (and indeed that is not feasible on most panel surveys). Moreover, some of the variance in responses will reflect respondent error as well as genuine shifts of attitude. Our assumption is that variability from these sources is random with respect to interview mode.

We carried out two main analyses. The first examines the small number (seven) of questions which appeared in all three panel waves; the second considers those which were not included in the BSA interview but appeared on both JUSST1 and JUSST2.

### 8.1. Questions appearing in all three waves

For the seven questions which appeared in all three waves of the panel there are a

Table 6.    Number of cases for difference combinations of modes for comparing stability over time

|  |  | BSA | JUSST1 | JUSST2 | Sample size |
|---|---|---|---|---|---|
| JUSST1 experiment areas: | Expt | Paper | Computer | Paper | 106 |
|  | Control | Paper | Paper | Paper | 98 |
| JUSST2 experiment areas: | Expt | Paper | Paper | Computer | 128 |
|  | Control | Paper | Paper | Paper | 132 |

number of possible comparisons which can be made. In those areas where computers were used in JUSST1 the sequence of interviews was paper-computer-paper (PCP) for one half of the sample and paper-paper-paper (PPP) for the other half (see Table 6). So in order to identify the possible effect of CAPI we can compare the stability of the two halves between the first and second waves, and between the second and third waves. If the stability of responses of the group who had computer interviews at wave two were lower than of those who were interviewed only by paper, we would conclude that CAPI is less reliable than PAPI. If on the other hand the responses of those who were interviewed by computer were more stable then CAPI would appear to be more reliable than PAPI. This would imply that the reliability of any time series would be improved by a change of mode.

For the areas where the computers were used in JUSST2 we can compare the level of stability between the second and third waves according to the mode used in the third wave. Again we are looking to see whether or not the stability of responses of those who were interviewed by computer is higher or lower than of those who were not. For this group a benchmark is also provided by the level of agreement between the first two waves which were both conducted by paper. Table 6 summarises the combination of modes at each stage and the numbers of cases available for analysis.

The seven questions included in all three waves had different numbers of response options (usually two or three and a "don't know" option) and were not intended to be ordinal measures. Levels of inter-wave stability are therefore shown in terms of the percentage of respondents giving the same answer at both waves (Table 7).

The two columns of the paper show the percentage of respondents who give the same answer at waves one and two (BSA and JUSST1) of the panel and at waves two and three (JUSST1 and JUSST2, respectively). This information is presented for each question in two sets of paired rows. In the first two rows we give the results for those respondents who lived in areas where some of the interviews at JUSST1 were conducted by computer. In the first of these rows the percentage agreement between waves is given for those respondents who were interviewed by computer on JUSST1; in the second row information is given for those interviewed by paper and pencil. In the second pair of rows we then give the same information for those respondents living in those areas where computer interviews were undertaken on JUSST2.

For example, if we look at responses to the first question, about British membership of the European Community, we see that 80% of respondents interviewed by paper gave the same answer in both the BSA and JUSST1 interviews compared with 84% of

*Table 7.    Percentage agreement over time for questions included in all three interviews*

| Question | Mode combination | | BSA/ JUSST1 | | JUSST1/ JUSST2 | |
|---|---|---|---|---|---|---|
| Membership of EC | JUSST 1 areas: | PCP | 84 | (PC) | 90 | (CP) |
| | | PPP | 80 | (PP) | 90 | (PP) |
| | JUSST 2 areas: | PPC | 78 | (PP) | 91 | (PC) |
| | | PPP | 84 | (PP) | 95 | (PP) |
| Membership of NATO | JUSST 1 areas: | PCP | 96 | (PC) | 99 | (CP) |
| | | PPP | 90 | (PP) | 95 | (PP) |
| | JUSST 2 areas: | PPC | 93 | (PP) | 91 | (PC) |
| | | PPP | 93 | (PP) | 95 | (PP) |
| Links with W Europe vs U.S.A. | JUSST 1 areas: | PCP | 58 | (PC) | 67 | (CP) |
| | | PPP | 68 | (PP) | 68 | (PP) |
| | JUSST 2 areas: | PPC | 66 | (PP) | 72 | (PC) |
| | | PPP | 70 | (PP) | 71 | (PP) |
| Siting of U.S. missiles | JUSST 1 areas: | PCP | 82 | (PC) | 83 | (CP) |
| | | PPP | 83 | (PP) | 89 | (PP) |
| | JUSST 2 areas: | PPC | 84 | (PP) | 87 | (PC) |
| | | PPP | 79 | (PP) | 87 | (PP) |
| Britain having own missiles | JUSST 1 areas: | PCP | 68 | (PC) | 73 | (CP) |
| | | PPP | 79 | (PP) | 82 | (PP) |
| | JUSST 2 areas: | PPC | 78 | (PP) | 79 | (PC) |
| | | PPP | 80 | (PP) | 76 | (PP) |
| Unilateral disarmament policy | JUSST 1 areas: | PCP | 76 | (PC) | 81 | (CP) |
| | | PPP | 72 | (PP) | 79 | (PP) |
| | JUSST 2 areas: | PPC | 79 | (PP) | 74 | (PC) |
| | | PPP | 76 | (PP) | 82 | (PP) |
| U.S. and Russia threat to peace | JUSST 1 areas: | PCP | 61 | (PC) | 60 | (CP) |
| | | PPP | 63 | (PP) | 64 | (PP) |
| | JUSST 2 areas: | PPC | 58 | (PP) | 69 | (PC) |
| | | PPP | 61 | (PP) | 62 | (PP) |

respondents who were interviewed by computer at JUSST1. Both sets of respondents had a 90% agreement rate between JUSST1 and JUSST2. These figures suggest that a change of mode did not affect the stability of responses. Turning to the results for the JUSST2 areas for this question we see that agreement between BSA and JUSST1 varied from 78% to 84% for the two groups of respondents, despite both groups having had paper interviews on both these waves. This suggests that there can be a fair amount of variation in the stability of responses without any change in mode of interview. This difference is indeed greater than that for JUSST1 and JUSST2 (91% for those who had a computer interview at JUSST2 and 95% for those who had a further paper interview) when there actually was a difference in the mode of interview.

*Table 8. Comparison of JUSST1 – JUSST2 correlations*

| Topic | Number of comparisons | Diff mode $r >$ same mode $r$ | Diff mode $r <$ same mode $r$ |
|---|---|---|---|
| Feelings about Britain | 18 | 12 | 5 |
| National symbols | 18 | 10 | 8 |
| Risks to environment | 14 | 8 | 4 |
| Feelings about Parties | 8 | 6 | 1 |
| Concerns of Parties | 24 | 13 | 11 |
| All comparisons | 82 | 49 | 29 |

The same conclusions emerge from most of the other comparisons we can make. The stability of responses of those interviewed on one occasion by computer and once by paper is for the most part virtually the same for those who were interviewed by paper both times. Where there are differences they are not in a consistent direction and can be accounted for as chance fluctuations. This analysis thus provides no evidence to support a claim that a change of mode increases error variance in the measurement of change.

## 8.2. Comparison between JUSST1 and JUSST2

For our second analysis, of questions asked in both JUSST1 and JUSST2 but not the BSA, we have far more items for comparison – 41 in all. As these items were all three or five point ordinal scales, we can measure reliability over time by calculating the correlations between the responses obtained on the two occasions. ("Don't know" responses were recoded to the central value; an alternative procedure of excluding them did not materially affect the results.) The correlation measures the extent to which respondents appeared to maintain their attitudinal position on the two occasions relative to the sample mean – which as we have already shown is unaffected by a change of mode. Again we are interested in whether or not the reliability of responses of those who were interviewed once by computer is higher or lower than those who were only interviewed by paper.

Because the sample sizes are too small to test differences in the inter-wave correlations of each individual item, we have classified each of the 82 relevant comparisons (two for each item) according to whether the correlation between responses of those interviewed only by paper was higher or lower than for those interviewed once by computer and once by paper. In Table 8 we have aggregated the results for each major question topic.

For four of the comparisons there was no difference between CAPI and PAPI respondents. Of the remaining 78, 49 showed a higher reliability amongst those who had been interviewed once by computer while only 29 showed a higher reliability amongst those who were only interviewed by paper. Using a two-tailed sign test as a simple test of the hypothesis that mode of interview has no effect on reliability gives a $z$-value of 2.26 which has a probability value of 0.02. Thus there is some suggestion in these results that if anything CAPI is a slightly more reliable method than PAPI. Taken together the two analyses suggest that a switch from PAPI to CAPI would not lead to greater instability of responses over time and might possibly lead to greater stability.

## 9. Summary and Conclusions

The interviewers reacted well to the use of computers for interviewing and handled their assignments without major problems.

There was some indication that rapport with respondents was not as good as for PAPI interviews and that for non-typists entering text was a problem. However these problems were less acute on the ESM study than the JUSST experiments, indicating that the problems are probably a function of lack of experience with CAPI and as such are likely to lessen if not disappear with time.

The study showed no evidence that a change to CAPI would affect willingness to participate on either a cross-sectional or a panel survey. Respondents accepted the computers with little comment, the majority expressing no preference between PAPI and CAPI. Taken with the lack of difference in response rates, this confirms the acceptability to respondents of laptop computers for interviewing.

The CAPI interviews on both JUSST experiments took longer than the paper interviews. However, for all the interviewers this was their first experience of using laptop computers. On the ESM study the majority of the interviewers had previously used the computers and this study showed no overall difference in length of interview between the two modes. The classification section, which required interviewers to enter text describing occupations, did however continue to take longer using computers. This was balanced by the complicated scale questions which, possibly because of a slight difference in their administration, took less time by computer. Overall these results suggest that once interviewers become experienced in using the computers, CAPI interviews are unlikely to take longer than PAPI interviews and are therefore unlikely to lead to an increase in the cost of interviewing.

The questionnaires used for all the interviews were fairly straightforward meaning that there was not very much missing

information for factual questions. Even so, there was some evidence that computer interviews produced less missing data than paper interviews.

Turning to attitude questions, CAPI and PAPI respondents showed no difference in propensity to choose the "don't know" option or to give any kind of neutral response. Nor was there any difference between them in the mean response to scale items. However, there was some evidence that CAPI respondents were more likely than PAPI respondents to choose the extreme categories of eleven or five point scales. It is not obvious at all why this should be. Evidence from other CAPI studies mentioned earlier suggests that greater reporting of sensitive behaviour arises because respondents think of CAPI as a more confidential mode than PAPI. Perhaps the same reasoning encourages respondents to be more willing to report what might be regarded as extreme views. Alternatively the use of CAPI might encourage respondents to give greater consideration to the differences between responses in making their choice, perhaps because the presence of a computer emphasises the seriousness of the exercise. With increasing familiarity with computers these considerations may however be ephemeral in which case the apparent higher incidence of extreme responses may disappear.

Using the panel design of the JUSST study to examine stability of responses over time provided no evidence that the introduction of CAPI would have an adverse effect on the reliability of questions; indeed there was some indication that it might lead to improved reliability.

The general conclusion of the study is that a change to computer assisted interviewing for attitude surveys is unlikely to have any major adverse effects on data quality, although some further investigation

of the apparent tendency to use extreme response options may be desirable.

## 10. References

Baker, R.P. (1990). What We Know About CAPI: Its Advantages and Disadvantages. Paper presented at the Annual Conference of the American Association for Public Opinion Research, Lancaster, PA, May 17–20.

Baker, R.P. and Bradburn, N.M. (1991). CAPI: Impacts on Data Quality and Survey Costs. Proceedings of the 1991 Public Health Conference on Records and Statistics, U.S. National Center for Health Statistics, 459–464.

van Bastelaer, A., Kerssemakers, F., and Sikkel, D. (1988). Data Collections with Hand-held Computers: Contributions to Questionnaire Design. Journal of Official Statistics, 4, 141–154.

Blom, E. (1990). Computer Assisted Data Collection Production Test on the Labour Force Surveys August 1989–January 1990. R&D Report, Statistics Sweden.

Bradburn, N.M., Frankel, M.R., Baker, R.P., and Pergamit, M.R. (1991). A Comparison of Computer Assisted Personal Interviews (CAPI) with Paper-and-Pencil Interviews (PAPI) in the National Longitudinal Survey of Youth. Paper presented at the Annual Conference of the American Association for Public Opinion Research, Phoenix, Arizona, May 16–19.

Catlin, G. and Ingram, S. (1988). The Effects of CATI on Costs and Data Quality: A Comparison of CATI and Paper Methods in Centralized Interviewing. In Telephone Survey Methodology, ed. R.M. Groves et al. New York: John Wiley and Sons, 437–450.

Couper, M., Groves, R.M., and Kosary, C. (1989). Methodological Issues in CAPI. Proceedings of the Section on Survey Research Methods of the American Statistical Association, 349–354.

Edwards, B., Edwards, S., Gay, N., and Sperry, S. (1992). CAPI on the Medicare Current Beneficiary Survey: A Report on Round 1. Paper presented at the 1992 Conference of the American Association for Public Opinion Research, May 18, 1992.

Groves, R.M. and Mathiowetz, N.A. (1984). Computer Assisted Telephone Interviewing: Effects on Interviewers and Respondents. Public Opinion Quarterly, 48, 356–369.

Groves, R.M. and Nicholls II, W.L. (1986). The Status of Computer Assisted Telephone Interviewing: Part II – Data Quality Issues. Journal of Official Statistics, 2, 117–134.

Manners, T. (1987). The First Laptop Computer Fieldwork Project in SSD. Survey Methodology Bulletin 21.

Sebestik, J.H., Zelon, H., DeWitt, D., O'Reilly, J.R., and McGowan, K. (1988). Initial Experiences with CAPI. Proceedings of the Fourth Annual Research Conference. Washington, D.C.: U.S. Bureau of the Census.

Sikkel, D. (1988). Quality Aspects of Statistical Data Collection. Sociometric Research Foundation, Amsterdam.

# Appendix

### Details of Design of the Experiments

The JUSST surveys were conducted amongst people who had originally responded to the 1989 British Social Attitudes (BSA) survey. The original BSA sample was a nationally representative sample of individuals selected from 152 areas (polling districts) of Great Britain. The respondents in one half of these areas were selected for inclusion in the JUSST

survey except that respondents who at the end of the original BSA interview had indicated their unwillingness to be re-interviewed were excluded. For practical reasons it proved impossible to include any areas in Greater London areas and interviews eventually took place in 67 areas. A total of 1192 BSA respondents were approached. The first wave of interviews (JUSST1) took place between November 1989 and January 1990, some four to nine months after the original BSA interview. The second wave (JUSST2) took place between May and July 1990, about a year after the original BSA interview. The overall response rates for the two surveys were 73% and 80%. This means that those successfully interviewed at JUSST2 comprised 58% of those originally selected for inclusion.

The laptop experiments were undertaken in 20 of these 67 areas. As indicated in the main text, 10 laptop computers were available to be shared among 20 interviewers. As each area was covered by one interviewer in the JUSST1 survey, this meant that 20 areas could be included. Within each area addresses were allocated randomly to CAPI and PAPI modes, 197 to CAPI and 200 to PAPI. The selection of areas was, however, non-random but consisted of those which contained the largest number of BSA respondents who had agreed to be re-interviewed. The aim was to maximise the sample size of the experiment. There is of course a danger that respondents living in such areas are more likely to be willing to be interviewed by CAPI than respondents elsewhere. This however is not a feature of the design of the JUSST2 or ESM samples.

For the JUSST2 survey there were insufficient numbers of respondents in each area to assign just one area to each interviewer and produce an adequately

sized sample; rather they were grouped together and each interviewer covered two or three areas. The selection of areas was largely determined by whether a sensible grouping of areas could be made with sufficient cases to be covered by one interviewer. In total 30 areas were included in the experiment, covered by 20 interviewers. It was intended that a different group of interviewers would use laptop computers for the JUSST2 interviews, thus maximising the number of interviewers who acquired experience of computer assisted interviewing. However, for practical reasons, two interviewers used computers for both waves of interviews.

The JUSST experiments suffer from the disadvantage of not starting with a fresh sample of respondents who have not been interviewed previously with paper questionnaires. However, the panel design offers some significant advantages. Information about respondents collected during the BSA interviews was available enabling differences in the characteristics of respondents and non-respondents according to mode of interview to be examined. In addition the stability of responses over time could be examined according to mode of interview.

The ESM study started with a fresh sample of 640 named electors selected from the 1990 Electoral Registers. The addresses were clustered in 20 areas (polling districts), with one interviewer per area each of whom was assigned 32 names. Within each area half the names were randomly allocated for a laptop interview and half for a paper and pencil interview. This design controls as much as possible for both interviewer and area effects since the CAPI and PAPI cases were split equally between interviewers within area.

As with the JUSST experiments, half the interviewers carried out the computer interviews first while the others started with their paper interviews. Half way through the

computers were passed to the other group of interviewers, allowing 20 interviewers to manage with only 10 computers. Seventeen of the 20 interviewers had worked on the JUSST experiments and therefore had previous experience of CAPI.

The questionnaire was as far as possible the same for both CAPI and PAPI, and was designed to be similar to that used on a real British Election Survey (BES), given that a General Election had not just taken place. All of the questions which form the main BES time series were included, as were many of the other questions included on the 1987 BES. Interviewing took place in October to December 1990. The overall response rate was 52%.

The CAPI software used for all the experiments was Blaise, developed by The Netherlands's Central Bureau of Statistics. The Blaise questionnaires were programmed to be as similar as possible to the paper versions of the questionnaires.