

Data Collection With Hand-held Computers: Contributions to Questionnaire Design

Alois van Bastelaer,¹ Frans Kerssemakers,¹ and Dirk Sikkels²

Abstract: The newly designed Netherlands Labour Force Survey is conducted with hand-held computers on a continuous basis from January 1987. In March 1986 hand-held computers were tested in a pilot study; over 1 400 respondents from 700 households were interviewed. The test confirmed earlier findings that hand-held computers are accepted without any problems by interviewers as well as interviewees. Consistency checks were specified in some parts of the questionnaire. Inconsistencies had to be corrected by the

interviewer. The quality of the questionnaire can be assessed by observing the interviewer's corrections and their paging backwards in the questionnaire (these manipulations were recorded by the computer). Inconsistencies remaining in the data when the interview was completed also suggest how the questionnaire can be improved.

Key words: Data editing; questionnaire design; CAPI; CATI; survey research.

1. Introduction

In March 1986 the Netherlands Central Bureau of Statistics conducted a pilot study to test a newly developed questionnaire for the Continuous Labour Force Survey with a hand-held computer (HHC). This experi-

ment can be considered from two viewpoints. First, it was the logical continuation of two earlier experiments, described in Bemelmans-Spork and Sikkels (1985a, 1985b), where HHCs were tested in the Price Survey and the Consumer Expectations Survey. Second, the experiment was a preparation for the Continuous Labour Force Survey which started in 1987. This survey aims at measuring labour market flows. The monthly sample size is 10 000 addresses.

The questionnaire in the pilot study consisted of two distinct parts. The first part on household composition was more structured than the household section in common paper-and-pencil questionnaires. Because the "head of household" concept had to be

¹ Department for Statistics of Employment and Wages, Netherlands Central Bureau of Statistics, 6401 CZ Heerlen, The Netherlands.

² Research International Nederland B.Y.

The views expressed in this paper are those of the authors and do not necessarily reflect the policies of the Netherlands Central Bureau of Statistics.

Acknowledgment: The authors are grateful to the editor and the referees for their comments and suggestions which improved this paper.

avoided, a sequence of questions was designed to establish the relations between the household members. In the second part, all household members age 15 or older were interviewed about their current labour market statuses and (retrospectively) about their statuses in the past 12 months; no more than three jobs within the past 12 months could be described. These labour market histories are reconstructed by starting from the current labour market status and then going back in time. Many consistency checks were included in the questionnaire on household composition; consistency checks were not yet included for the labour market questions.

In this paper we describe some results of the pilot study, focusing on interviewer behaviour and questionnaire design. Section 2 sketches a short history of computer assisted data collection. In Section 3, the pilot study and the software for the data collection are described. Section 4 deals with the acceptance of the HHC by interviewers and respondents. In Section 5 the inconsistencies in the data are discussed and conclusions are drawn from an analysis of these inconsistencies. An example of the detection of an ill-defined concept from an analysis of inconsistencies is described in Section 6. Section 7 deals with paging backwards in the questionnaire. Section 8 concludes.

2. History

The history of computer assisted interviewing goes back to 1971, when AT&T sponsored a CATI (Computer Assisted Telephone Interviewing) survey to measure customer evaluations of telephone services (Nicholls and Groves (1986)). Since then, CATI has developed rapidly throughout the world. Today it is a common tool for commercial market research, governmental statistics, and scientific purposes. Despite its 17 year history, little research has been done on the methodology and data quality of CATI. In

market research it seems that the main consideration for the introduction of CATI is cutting costs compared with face-to-face interviews, which usually implies that a CATI system must be flexible and very simple to use. Once such demands are fulfilled the users are satisfied and feel no need for further research. This experience was confirmed by Groves and Nicholls (1986), who in their comprehensive paper, stated that "... there is little reliable empirical evidence that CATI affects data quality. This absence is especially noteworthy in the context of the frequent expectation of data quality improvement of CATI."

CAPI (Computer Assisted Personal Interviewing, interviewing with HHCs) began its development when computers that were small and powerful enough appeared on the market. As observed by Shanks (1983) and Shanks and Tortora (1985), CAPI questionnaires have much in common with CATI questionnaires. CAPI and CATI are in essence a sequence of displays which depend on previous responses. This dependence may be complicated. In this way the construction of a questionnaire closely resembles the development of a computer program. House (1985) and Jabine (1985) observe that the design and documentation of a CATI questionnaire requires the same approach as ordinary computer software. A significant difference between CAPI and CATI is that CAPI does not impose extra limitations on the complexity of the questions and response categories and that CAPI allows for visual aids such as display cards.

The first test of HHCs that we are aware of was carried out by Statistics Sweden, in their Price Survey, using a pocket calculator sized computer (Danielsson and Maarstad (1982)). Later the Netherlands Central Bureau of Statistics started experimenting with HHCs. In Holland too the Price Survey was used for tests with a larger HHC, the size

of a telephone directory (Bemelmans-Spork and Sikkel (1985a)). Both experiments were successful in that they proved that the use of HHCs was possible, although the hardware needed improvement. From there on the strategies of Holland and Sweden diverged. Statistics Sweden formulated demands for an HHC that corresponded to its needs and found a manufacturer to develop this computer (Lyberg (1985)). The Netherlands Central Bureau of Statistics continued experiments with HHCs that had been developed for a more general market. Around Christmas 1984 a trial was carried out with a simple household survey, the Consumer Expectations Survey, see Bemelmans-Spork and Sikkel (1985b). In this experiment one group was interviewed with HHCs, and another with paper-and-pencil. Both groups consisted of about 175 respondents. The comparison showed no differences in unit nonresponse or item nonresponse for sensitive questions. Nor were there significant differences between the groups in the frequency distributions of the relevant variables. The first CAPI survey for production purposes was conducted in 1983 in a rather different context. Bus passengers in Durham, U.K., were asked simple questions to which the answers were entered into a computer that was a little larger than a pocket calculator. Hamilton (1985) reports that this survey was satisfactory with respect to the fieldwork as well as to economic aspects; the investments paid for themselves within a year.

3. The Pilot Study and the Software for Data Collection

Since the findings from the previous experiment were felt to be satisfactory, the pilot study of the Continuous Labour Force Survey was not designed to measure subtle differences between two different groups.

Instead more quantitative evidence was sought about the acceptance of CAPI by interviewers and respondents and about any problems connected with the use of hand-held computers for data collection. A second objective was to obtain information on the data resulting from interviews with a complex questionnaire. The reduction of measurement errors was also analysed by comparing earlier false entries with the final correct entries in the same interview. In this way the pilot study was highly useful for the design of the final questionnaire to be used from 1987.

During the last three weeks of March 1986, 23 interviewers visited 1 224 addresses in ten municipalities. They had received a training of one full day and two half days. The interviewers used an HHC, type NEC PC-8201A with two memory banks of 32Kb RAM each. One bank contained the questionnaire and answers. The BASIC-program QUEST2 that interprets the questionnaire and a module for data communication were loaded on the other bank. QUEST handles the display of the question-texts and the response categories; it handles data entry and data storage in a compressed form which is to be decoded into fixed field records after the data are received by the central host computer. It checks the specified data inconsistencies and it manages the branching and skipping. It is the hardware, not the software that limits the number of questions. The definition of a question consists of four sections: in Section 1 the question type (precoded, numeric, open ended, etc.) was defined and the question text entered; Section 2 contained the response categories or the response range; Section 3 specified edit checks; and Section 4 contained the codes for branching and skipping. These codes can depend on the logical or numerical operations of previous entries. Text strings are often defined as variables for repeated use, also depending on previous answers. Besides saving memory, this allows

the clear-cut phrasing of questions and response categories. The detection of an inconsistency with prior answers causes the relevant questions to be displayed one after the other for confirmation or correction until the answers are reconciled.

The interviewers had a number of special programmable keys at their disposal in addition to the regular data-entry keys: return to the immediately preceding question (programmable key 1), return to any previous question (shift + programmable key 1), confirm a previous entry after having returned to a previous question (key 2), add remarks (shift + key 2), no choice in a multiple choice question (key 3), return to the current question while displaying all intermediate questions that have already been answered (shift + key 3), do not know (key 4), immediate return to the current question skipping intermediate questions (shift + key 4), refusal (key 5), refusal further cooperation with the interview (shift + key 5), consult table with household data (cursor right) and display question number and bytes still free (cursor left). Some programmable keys were confusing (e.g., return to the current question with and without displaying intermediate questions) and some were redundant (confirming an entry with key 2 whereas the enter/return key was used for the current question). In the revision of QUEST this user interface was redefined.

Each HHC was programmed to automatically phone the central host-computer at the office at a specified time during the night. These times were different for each HHC and distributed uniformly throughout the night. The data transmission took a few minutes, and the quality of the data transmission was examined through check sums. If the data transmission was successful, the data were released from the HHC so that the HHC could be used for new interviews. Following some problems during the first few

days, the communication ran smoothly for the remainder of the test. This led to the conclusion that communication by phone works well (at least given the quality of the Dutch telephone network, which is fairly high). The value of tailor-made software for communication, however, should not be underestimated.

Due to the short fieldwork period, the total response rate was rather low: 56 %, i.e., 1 407 persons of age 15 or older participated.

4. Acceptance by Interviewers and Respondents

Danielsson and Maarstad (1982) and Bemelmans-Spork and Sikkell (1985a, 1985b) gave the impression that the HHCs were readily accepted by interviewers and informants. Due to the relatively large sample in our experiment, we now are able to confirm these impressions with more solid results from two evaluation forms filled in by the interviewers. One form was filled in for each responding household, the other after each week of interviewing.

The following questions were answered by the interviewer following each interview.

- Did the respondent's attitude change noticeably when you showed him or her the hand-held computer? (Table 1)
- Did the respondent comment on the use of the hand-held computer?
- If so, how?
- Did you feel that the hand-held computer caused any inconvenience for the respondent when answering the questions? (Table 2)
- Did the respondent inquire about the data processing or about confidentiality?

In addition to questions about the structure and content of the questionnaire the interviewers had to answer the following questions every week.

- Did typing texts for questions on economic activity or occupation interfere more with the interview than writing the answers on a paper-and-pencil questionnaire? Why?
- Did the hand-held computer refuse any answers (e.g., was the message “are you sure?” displayed)?
- Did you have to return to previous questions to correct mistakes?
- Were there any problems with returning to previous questions?
- Were there any problems with the hand-held computer?
- Could you enter answers that you already knew (e.g., on the composition of a household) fast enough?
- Did the hand-held computer cause any

- delay when you wanted to ask a new question or enter the answers?
- Which questions caused problems?
- Were there any problems with the modem?
- Was the readability of the screen sufficient?
- Do you prefer working with a hand-held computer or a paper-and-pencil questionnaire? (Table 3)

First we shall give some results of the forms that were filled in per household. The interviewers were instructed to show the HHCs only after the respondent had agreed to the interview. In the evaluation form there was a question about the respondent’s reaction, see Table 1.

Table 1. Respondents’ first reaction to the HHC

	1986		1984	
	Absolute	%	Absolute	%
No reaction	667	92.4	113	65.3
Positive reaction (e.g., interested)	36	5.0	34	19.7
Neutral reaction (e.g., surprised)	9	1.2	21	12.1
Negative reaction (e.g., suspicious)	10	1.4	5	2.9
Total	722	100	173	100

Here the results are compared with the previous experiment in 1984. In neither was there any extra nonresponse due to the HHC. The negative reactions were almost negligible, and the following are typical examples: “automation strikes again” or “can we be recorded?” Examples of favourable reactions are: “very interesting,” “called her husband because she thought it was fantastic,” “how nice, is that a tape-recorder (or typewriter)?” and “will we be on television?” Most surprising, however, was the increasing number of respondents who did not react at all. This suggests that there is a growing acceptance of the computer as a common

tool. A second question was: “Did you feel that the hand-held computer caused any inconvenience for the respondent when answering the questions?” The answers are displayed in Table 2.

Table 2. Did the HHC cause any inconvenience for the respondent?

	Absolute	%
No	701	97
A little	20	3
Very much so	1	0.1
Total	722	100

The vast majority of interviews presented no problems. Another question concerned confidentiality. About 100 respondents asked about this but most of their comments would also have been valid for paper-and-pencil interviews. They were the normal questions on the method of data processing, data protection, the retainment of anonymity, and the possibility or probability of linking interview responses to names and addresses. Other comments were more specifically related to the use of the HHC. Respondents expressed distrust of the way the answers were recorded, approval that the answers could not be changed once they had been entered in the HHC, and curiosity over how the data were transmitted to the office (a frequent question). Most respondents could be convinced that CAPI guaranteed confidentiality better than paper-and-pencil questionnaires because the answers are stored in a compressed form separate from the questions and because the answers are recoded, these new codes are not identical representations of the (alpha)numerical entries.

In the weekly evaluation forms, the interviewers were questioned about several aspects of CAPI. In the 1984 test it appeared that the quality of the light in the respondents' houses affected the readability of the HHC screens. In the current test there was a specific question about readability; about one in five interviewers complained about the poor readability of the screen.

In the first week, six interviewers felt that

using the keyboard to enter text strings interfered more with the interview than writing answers on paper questionnaires. In the second and third weeks only two or three interviewers retained this opinion. The interviewers gradually became more accustomed to the keyboard for data entry.

During the first week almost half the interviewers complained about the slow speed of the program, especially when data had to be entered that did not need to be asked or were already known. This combined with complex skipping patterns and consistency checks slowed down the reponse time of the program to about two seconds. In the second and third weeks fewer interviewers complained about the speed; here too they grew familiar with the hand-held computer and the questionnaire. Meanwhile better hardware and other software (a Pascal instead of Basic program) have considerably improved the performance of an interview with the hand-held computer.

The hand-held computer did fail now and then, mostly because of lack of electrical power or because of program bugs or disconnected chips.

The reported problems may suggest that the interviewers had a bad attitude towards CAPI. This, however, was not the case as appears from a general comparison of hand-held computer and paper-and-pencil questionnaire (Table 3). Here we distinguish between the first and the third week of the experiment.

Table 3. Comparison of CAPI and paper-and-pencil

	Week 1		Week 3	
	Absolute	%	Absolute	%
CAPI better	11	52	10	71
Neutral	4	19	1	7
Paper-and-pencil better	6	29	3	21
Total	21	100	14	100

The percentage of interviewers who preferred CAPI to paper-and-pencil increased from 52 in the first week to 71 by the third week (the different totals of interviewers are caused by the fact that not every interviewer participated every week). Some favourable comments by the interviewers: “after two weeks better, after three weeks good, no longer uncertain,” “much more convenient, no more paperwork,” and one negative: “the interview is less natural.”

5. Exploring the Questionnaire Design

5.1. Introduction

For each interviewer, the HHC also recorded important information about the flow of the interview, for example, returns to previous questions to consult the answers or to correct errors and answers prior to corrections. These data describe the error-checking and may point out ill-defined or poorly understood concepts in the questionnaire.

Error checks specify that the answer to question R must be in range Y if the answer to a previous question Q lies in range X. If such a condition is not satisfied, the HHC queries “are you sure?” and asks question R again. If the interviewer confirms question R, then the previous question Q is asked again. Only after confirming this question and once again confirming question R is the

interviewer permitted to enter an inconsistency.

5.2. True value: the number of household members

A true value can be assessed if two different questions with a common content are asked and if the answer to the second question is redundant. The latter information can then be used as a check to the former answer and vice versa to determine the true value. A simple example from the household questionnaire illustrates how error checks can be defined on two questions with a common content. These questions are “how many household members are there?” and “is there anyone else in the household?” The latter question is asked repeatedly after the data for each person are completed. The answers must be consistent with the previously given number of household members.

Another error check consists of two questions about the number of household members and the household composition. A couple with children consists of at least three members, a couple with children and others of at least four. The purpose of the questions about the number of household members and the household composition is to give the interviewer a preliminary overview before she proceeds with the questions for each household member. The response categories

Table 4. Broad composition of the household

		Absolute	%
*	Single household	193	28
a.	(Un)married couple alone	180	26
b.	(Un)married couple + child(ren)	264	38
c.	(Un)married couple + child(ren) + other(s)	6	1
d.	(Un)married couple + other(s)	1	0.1
e.	Single parent + child(ren)	30	4
f.	Single parent + child(ren) + other(s)	2	0.3
g.	Other (household core not: (un)married couple or single parent)	10	2
Total number of households		686	100

and the corresponding frequency distribution of the second question are given in Table 4.

In this question the concept of the household core instead of the head of the household was central. In cases *a* to *d* this consisted of the couple, for *e* and *f* it was the single parent. The interviewers were instructed first to enter the data of the household core, then those of the children (of one or both members of the household core) and finally those of the others. The family relations were always a relative of a member of the household core, where possible the respondent.

The conflicts that actually occurred and

the reactions of the respondent are displayed in Table 5. It appears that 46 conflicts were detected, 2 of which remained in the final data (confirmed twice). In one case the household was reported to consist of 4 members, after which the data of only 2 persons were entered. In the second case 33 members were reported followed by data on 3 persons. Altogether 36 conflicts involved the related questions of the number of household members and the question of whether there was still another household member. Of these, 33 were immediately resolved by changing the answer to the latter question.

Table 5. Conflicts with the number of household members

Current question	Conflicting answer to current question		Confirmation current question	Confirmation previous question	Total number of conflicts
Another person in household after	yes	no			
person 1	6	17	0	0	23
person 2	3	2	1	1	5
person 3	0	4	1	1	4
person 4	3	0	0	0	3
person 6	1	0	1	0	1
	Couple without children	Couple with children			
Composition household	9	1	0	0	10
Total			3	2	46

5.3. True value: the household composition

Another illustration applies to error checks on the household composition. If there are children in the household according to the preliminary household composition reported, then the third (for single parent families, the second) person should be a child. Actually the questionnaire was programmed in such a way that the interviewer had only to confirm that the next person was indeed a child. If this was denied, the conflict had to

be solved. Also if there were no “others” then all persons not belonging to the household core must have been children. If categories *a* to *d* of the household composition (Table 4) were entered, then it could not be denied later without a conflict that there was a couple in the household, et cetera. The various categories of the household composition resulted in equally varied patterns of conflict, as described in Table 6. The question on the couple’s marital status was asked

Table 6. Conflicts with the household composition

Current question	Conflicting answer to current question			Confirmation current question	Confirmation previous question	Total number of conflicts
	married couple	unmarried couple	no couple			
Unmarried or married couple	0	0	3	0	0	3
Child of person 1?	yes	no				
person 2	0	3		1	1	3
person 3	0	4		3	2	4
person 4	0	2		2	1	2
Relation to person 1	child	other				
person 5	2	0		1	0	2
Total				7	4	14

only when the household composition indicated that there was a couple. The answer "no couple," therefore, always created a conflict which in the pilot study was always solved. All remaining inconsistencies were between the reporting of children in the household composition and the not reporting of these children later on in the household box.

5.4. Other conflicts

The number of other conflicts was so small that we do not present them in a table. These conflicts concern the network of relations established within the household box. In five cases it appeared that the relation of person B to person A was in conflict with the marital status of person A. This relation was corrected three times; the marital status was corrected once; one inconsistency was confirmed. This concerned a remarried widower who still wanted to be considered a widower. Finally two conflicts appeared between the marital status of person B and the question about a married or unmarried couple. In both cases the marital status of person B was changed to "married."

5.5. True value: the date of an event

The software can handle a variety of labour market histories with a maximum of three jobs in the past twelve months. One respondent may have had a single job for a few years already, another respondent may have changed jobs twice in the past twelve months while being unemployed between two jobs. These event histories were established in the questionnaire by introductory questions which determined the correct path through the questionnaire. The response categories of the introductory questions were three or six month periods. After the introductory questions, the specific dates of beginnings and terminations of jobs were asked. Obviously, these dates have to satisfy certain order relations. A new job should have started after a previous job started, et cetera. However, no error checks were specified. In the absence of error checks some inconsistencies remained (Table 7).

The inconsistencies of cases 3, 4, and 10 are violations of the order relations of the dates. They are caused mainly by respondents who usually do temporary work and therefore have a complicated labour market

Table 7. Inconsistencies of dates with the introductory questions and violations of chronology

General labour market history		Inconsistencies and violations	Specific labour market history
1.	B	start current job less than 1 year ago (4x)	C
2.	C	start current job more than 1 year ago (2x)	B
3.	C	start current job after date of interview (1x)	
4.	D	start last job coincides with end last job (5x)	
5.	D	end last job precedes start last job (1x)	
6.	F	start previous job more than 1 year ago (1x)	G
7.	G	end previous job coincides with start current job (1x)	
8.	H	start previous job more than 1 year ago (1x)	I
9.	J	start earliest job more than 1 year ago (1x)	K
10.	K	start previous job coincides with start current job (1x)	

For the meaning of the capitals indicating the type of history, see the Appendix.

history. Case 7 is not a violation of an order relation, but rather a false entry. The end of a previous job coincides with the start of a current job even though data are also provided on the intermediate period.

The inconsistencies of cases 1, 2, 6, 8, and 9 are of another type. Here the broad indication does not correspond with the specific dates.

Table 7 may give the impression that the number of inconsistencies is insignificant. This is true if compared with the total sample of 1 407 respondents. However, if there are, say, 6 inconsistent records among the 19 records with a single job in the last 12 months that started and ended within these 12 months (labour market history D), then error checks are necessary for more complex labour market histories.

5.6. Interpretation

In CAPI and CATI, consistency checks have two purposes. First, they serve to ensure that the answers are entered correctly. If any errors are detected by the HHC program, they can be rectified during the interview. Consequently, the resulting data set is error

free and may even be statistically processed directly. Second, with computer assisted interviewing, data consistency is essential because responses or series of responses lead to intricate branching and skipping patterns later in the interview. Interviewers are expected to perform better because the hand-held computer takes care of the routing of the questions and because questions are displayed one at a time, thus focusing the interviewers' attention on that particular question. From time to time, however, inconsistencies are inevitable. Our most striking example was the married man who insisted on being registered as a widower. The HHC must allow for such a conflict, but it makes the task of the questionnaire designer far from easy.

A study of the conflicts that occurred in the field can contribute to developing the methodology of questionnaire design. In this way the designer may develop an intuition about the effects of error checks. The results of our pilot study are insufficient to draw definite conclusions, but they are sufficiently suggestive to generate a few recommendations. These recommendations may be generally valid, i.e., independent of the interviewing mode.

If error checks are specified between two different questions with common content, a true value can be assessed. When a true value is assessed during the interview, the measurement error is reduced. The pilot study of the Labour Force Survey provides evidence about the assessment of a true value. Because error checks can be specified between two questions with common content, both questions should be retained in the questionnaire. Error checks are also effective when the answers to general and detailed questions have to be reconciled.

The number of corrections per interview may be interpreted as an indicator of the quality of a questionnaire. However, questionnaires on different topics cannot be compared since some topics are simply more complex than others. Hence a questionnaire – *ceteris paribus* – with the least number of corrections needed, is preferred.

If a survey organization wants to test a questionnaire, the software for computer assisted interviewing should be able to keep a record of all interviewer actions or, even better, to produce summary statistics on consistency checks.

6. Detecting Ill-defined Concepts

The occurrence of inconsistencies sometimes indicates an ill-defined or poorly understood concept. A surprising and important example is the parent-child relation in our definition of household. Many conflicts and most remaining inconsistencies had to do with this relation.

The concept “child” may have more than one interpretation in an interview and should therefore be well-defined. There are different reasons for confusion. The word “child” may be interpreted as the opposite of “grown up” or “married child.” This may also create confusion in determining the household core (e.g., a single parent household core implies

the presence of a child). Moreover there is the problem of adopted children or stepchildren, e.g., whether for a married couple a child of the second partner from a former marriage and thus of the household core according to the first question, will also be considered a child of the first partner to whom the relation is determined in the second question (by confirmation).

7. Detecting Deficiencies in a Questionnaire: Paging Backwards

A difference between computer assisted interviewing and paper-and-pencil interviewing is that the interviewer loses her overview to some extent because she can see only one screen at a time. With the HHC the interviewer could page backwards, question by question, using a programmable key. This option not only gives the interviewer an overview when necessary, but is also used to correct previous answers. In this way, paging backwards can be considered a form of error checking that was not specified by the designer of the questionnaire. The HHC kept a record of the interviewer’s backwards steps in the questionnaire.

Table 8. Number of steps needed to reach the desired question

Steps	Number	Percentage	Cumulative
1	664	85.46	85.46
2	67	8.62	94.08
3	21	2.70	96.79
4	8	1.03	97.82
5	6	0.77	98.59
6	1	0.13	98.72
8	4	0.51	99.23
9	2	0.26	99.49
11	1	0.13	99.62
12	1	0.13	99.74
17	1	0.13	99.87
29	1	0.13	100.00
Total	777	100.00	

Table 9. Questions that were referred back to more than 10 times

Question no.	Short description	Number of times back to	Number of times asked	Percent back
1	Result household: response/nonresponse?	12	686	1.75
2	Number of members of household?	14	686	2.05
107	Which member of the household answers questions?	17	1400	1.21
109	Activities of respondent (multiple choice)?	44	1407	3.13
111	Do you have a paid job?	12	787	1.52
112	Have you ever had a paid job?	11	733	1.50
121	Name of the company where you work?	15	598	2.51
122	Address of the company where you work?	18	598	3.01
124	In which department or place of the company do you work?	16	598	2.68
169	What is your occupation?	10	598	1.67
170	What are your main activities?	13	598	2.17
179	How many hours do you work (SWW ¹)?	27	281	9.61
180	How many hours do you work (no SWW ¹)?	18	306	5.88
181	Did you work longer last week?	22	598	3.68
184	Full time or part time?	11	598	1.50
213	Did you have another job before this one?	11	656	1.95
343	How long have you been looking for this (2nd) job?	10	77	12.99
372	How long have you been looking for a job?	13	60	21.67
423	Have you been looking for a job during the last months?	13	630	2.06
445	Sequence number address?	10	686	1.46
448	Number of households at this address?	14	686	2.05
459	Day of first visit?	18	686	2.62
460	Time of first visit?	14	686	2.05

¹ SWW = shorter working week.

The average number of times per household an interviewer paged backwards was 1.132. Table 8 shows how many steps the interviewers had to take to reach the desired question. In most cases one step was sufficient. In 67% of the cases paging backwards served to correct a previously given answer, and in 33% of the cases the previous answer was confirmed. This is consistent with findings in the CATI case reported in Groves and Nicholls (1986).

Since the “previous question” key could be used anywhere in the questionnaire, it is impossible to list all the 172 questions that were paged back to without reproducing a substantial part of the questionnaire. We therefore restrict ourselves to those questions to which the interviewers jumped back 10 times or more. In Table 9 these questions are de-

scribed briefly. An indication is also given of how many times the question was asked (obviously you cannot go back to a question that was not asked). Moreover, the percentage of “back jumps” is indicated. This leads to some remarkable results. The question that was most frequently jumped back to (109) has a relatively low “back jump” percentage because this question was asked to everyone. The percentage is much higher for questions 343 and 372 because only few respondents answered these questions.

A first conclusion of the analysis of paging backwards may be that the interviewers are concerned with the quality of their work. This is apparent from their frequent use of the “previous question” key. It was most often used to correct previously given answers. Given the preliminary status of the

questionnaire we shall restrict ourselves to only a few topics.

In question 109, the respondent was asked to report his activities in a multiple choice question (paid work, student, housewife et cetera). In question 110 the respondent had to make a single choice from the same alternatives (the most important activity). The fact that the interviewer paged back 44 times shows that this construction was poorly understood. In the questions determining the respondent's occupation the interviewers often paged back, indicating that this is a difficult subject. But on the other hand it is hard to improve such questions. Questions about working hours also cause problems especially when there is a shorter working week (questions 179–181). The retrospective questions 343 and 372 have the highest percentage of referrals to the “previous question.” Apparently it is difficult to answer questions about a complicated labour market history.

The relatively low number of referrals to the “previous questions” in the household box is remarkable, but it may be explained by the relative simplicity of the subject. An alternative explanation is the abundance of consistency checks in the household box. Detected inconsistencies mostly lead to correction (see Section 5), so that there will be less need for correction by paging back.

Consistency checking and paging back are probably not independent processes. Also, when certain questions are spontaneously corrected by the interviewer, these questions may be identified as error prone, or, from another perspective, questions where error checks are effective.

It may also be asked whether the number of times the interviewers paged back is a good indicator of the quality of the questionnaire. As with error checking, the answer is not a simple “yes.” The frequency of paging back is probably a good indicator of the difficulty of a questionnaire. Of course it is important to keep questionnaires as simple as possible,

but there are simple topics as well as difficult topics, such as occupation or economical activity, about which we have to gather information. Therefore, the preferred questionnaire is one in which the interviewers page back least often, given the complexity of the subject.

8. Conclusion

This paper had two goals. First, it aimed at presenting new findings about acceptance and appreciation of CAPI by interviewers and respondents. Acceptance by respondents is no problem; to them CAPI circumvents rather than arouses suspicion about confidentiality. Interviewers' attitudes are also favourable: the majority of the interviewers consider a hand-held computer more convenient than paper-and-pencil.

The second goal was to contribute to the development of a methodology for CAPI. The results suggest that error checks in CAPI are necessary. This is consistent with Tortora (1985) who compared the use of error checks for CATI and paper-and-pencil. Moreover, keeping a record of the interviewer's consistency checks and movement through the questionnaire is one step in evaluating the quality of the questionnaire.

9. References

- Bemelmans-Spork, M. and Sikkel, D. (1985a): Observation of Prices With Hand-held Computers. Statistical Journal of the United Nations Economic Commission for Europe, 3, (2).
- Bemelman-Spork, M. and Sikkel, D. (1985b): Data Collection With Hand-held Computers. Proceedings of the 45th session. International Statistical Institute, Book III, topic 18.3.

Appendix. The types of labour market histories and their distribution; having a job is indicated by x; not having a job by –.

Labour market history	n	1 year before interview			Date of interview
A	682	————	————	————	
B	536	xxxxxxx	xx		
C	63	————	————xxxxxxxxxxxxxxxxxxxxxxxx		
D	19	————	————xxxxxxxxxxxxxxxxxxxx		
E	44	–xxxxxxxxxxxxxxxxx	————		
F	6 [1] ¹	————	————xxxxxxx————xxxxxxxxxxxxx		
G	46 [35]	–xxxxxx	xxxxxxx————xxxxxxxxxxxxxxxxx		
H	2 [1]	————	————xxxxxxxxx————xxxxxxxxx		
I	3 [2]	–xxxxxx	xxxxxxxxxxxxxxxxx————xxxxxxxxx		
J	1 [1]	————	–xxxxxxxxx————xxxxxx————xxxxxxxxx		
K	4 [2]	–xxxxxx	xxxxxxxxx————xxxxxx————xxxxxxxxx		
L	1 [1]	————	–xxxxxxxxx————xxxxxx————xxxxxx		

¹ Two jobs may succeed each other immediately from labour market history F; this distinction is not made in the figure, i.e., a period indicated by “—” may be empty. The number in [...] indicates the number of persons with immediately succeeding jobs.

Danielsson, L. and Maarstad, P.A. (1982): Statistical Data Collection With Hand-held Computers – A Test in Consumer Price Index. Statistics Sweden, Örebro.

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

Hamilton, T.D. (1985): Hand-held Data Capture Devices. Paper presented at the meeting of the Study Group on Computers in Survey Analysis, London.

House, C.H. (1985): Questionnaire Design With Computer Assisted Telephone Interviewing. Journal of Official Statistics, 1 (2), pp. 209–220.

Jabine, T.B. (1985): Flow Charts – A Tool for Developing and Understanding Survey Questionnaires. Journal of Official Statistics, 1 (2), pp. 189–208.

Lyberg, L. (1985): Plans for Computer Assisted Data Collection at Statistics Sweden. Proceedings of the 45th session. Inter-

national Statistical Institute, Book III, topic 18.2.

Nicholls, W.L. and Groves, R.M. (1986): The Status of Computer-assisted Telephone Interviewing: Part 1 – Introduction and Impact on Cost and Timeliness of Survey Data. Journal of Official Statistics, 2 (2). pp. 93–115.

Shanks, J.M. (1983): The Current Status of Computer-assisted Telephone Interviewing: Recent Progress and Future Prospects. Sociological Methods & Research, 12 (2), pp. 119–142.

Shanks, J.M. and Tortora, R. (1985): Beyond CATI: Generalized and Distributed Systems for Computer assisted Surveys. Proceedings of the first annual research conference of the Bureau of the Census, U.S. Bureau of the Census, Washington D.C., pp. 358–377.

Tortora, R. (1985): CATI in an Agricultural Statistical Agency. Journal of Official Statistics, 1 (3), pp. 301–314.