Developments in Survey Methodology
Anuška Ferligoj and Andrej Mrvar (Editors)
Metodološki zvezki, 15, Ljubljana: FDV, 2000

Technology Effects: Interview Duration in CAPI and Paper and Pencil Surveys

Marek Fuchs[1], Mick P. Couper[2], and Sue Ellen Hansen[2]

Abstract

The transition from paper based personal interviews to computer assisted personal interviewing (CAPI) is already well underway. Much of the early research focused on operational issues and concerns about data quality differences between the methods. Attention is now being turned to more detailed assessments of specific features of the new data collection technology and its impact on the survey process. This paper deals with the question of relative administration of survey questions between paper and pencil and CAPI modes. Using data from a series of interviews (14 PAPI and 37 CAPI) using the National Health Interview Survey (NHIS) instrument conducted as part of usability testing of the instrument, a large number (over 2,200) of comparable items from the socio-demographic part of the instrument were subjected to detailed time and activity coding. These data allow us to examine reasons for time differences across the modes. Where items are comparable in terms of design across modes, we find that CAPI takes slightly longer than PAPI, largely due to the speed of typing versus writing. However, most of the time differences found can be attributed to differences of design between paper and pencil and CAPI, rather than as a result of the technology itself.[3]

[1] Faculty of History and Social Sciences, University of Eichstätt, Ostenstrasse 26, 85071 Eichstätt, Germany; Marek.Fuchs@ku-eichstaett.de

[2] Institute for Social Research, University of Michigan, Thompson Street, Ann Arbor, MI 48104, USA; MCouper@isr.umich.edu, SEHansen@isr.umich.edu

[3] The work reported here was financed by a research stipend of the Humboldt Foundation of Germany to the first author who stayed at the Survey Research Center, University of Michigan as a visiting scholar in 1996/97. The video tapes used for the secondary analysis were conducted under a cooperative agreement (principal investigator: James M. Lepkowsky) supported by the US National Center for Health Statistics.

1 Interview duration in a CAI environment

In surveys applying computer assisted interviewing (CAI), data entry and data editing take place in the interview situation. So far it is not clear what effect this additional task has in terms of interview duration. It might be that time savings during post-interview processing are outweighed by a longer duration of the interview itself. There are several studies dealing with this issue. However, these experiments show heterogeneous results. In this paper we review the existing findings on interview duration under CAI conditions and subsequently report results of a small scale experiment.

So far, there is little empirical evidence that the total survey duration is affected by the introduction of computer assisted interviewing. Even though the total project duration does not increase a shift from post-interview tasks toward pre-interview tasks was assumed. According to the larger amount of time necessary for preparation, development, and training, studies applying computer assisted interviewing are less time consuming at the back-end. As a result we observe time savings related to fielding and post interview processing. However, computer assisted interviewing is assumed to be especially time-effective for studies on a regular basis or panel studies in which preparation and development does not start from scratch every time a survey is fielded. It can be seen as part of this diminishing weight of post-interview processing, that with surveys applying computer assisted interviewing (CAI), data entry and data editing take place during the course of the interview. So far it is not clear what effect this additional task has in terms of interview duration.

What do we know about CAI interview duration? Early studies conducted in the 80s show that computer assisted interviewing takes longer than paper-based interviewing (Catlin and Ingram, 1988; Brikett, 1988; Harlow et al., 1985; Groves and Mathiowetz, 1984; Waterton and Duffy, 1984). The differences were mainly attributed to hardware specifics and software problems. Especially the fact that the early laptops used in these studies did not perform as fast as necessary was recognized as one of the factors contributing to the additional interview duration. Further studies found that interviewers could not get used to the new technology so that handling the hardware and software takes more time than working with a paper instrument. Almost 10 years ago Couper and Groves summarized these experiences: "whether this will change over time as interviewers become more familiar with the system or whether it is an inherent feature of CAPI is not yet clear" (Couper and Groves, 1989: 350).

In recent years, several studies were conducted concerning the differences between paper based personal interviews (PAPI) and computer assisted personal interviews (CAPI) in terms of interview duration. These studies show heterogeneous results:

- Baker and colleagues (Baker, 1992; Baker et al., 1994) conducted an experimental study with the National Longitudinal Study of Youth using similar instruments on paper and on a laptop computer. They found an average interview duration of 57 minutes for PAPI interviews and an average of 47 minutes for the CAPI condition. These differences increased according to the interviewers′ experience with the particular instrument.
- In another study, Lynn and Purdon (1994) compared the two technologies within a British attitude survey. They found clear evidence that computer assisted interviewing takes less time.
- Martin and colleagues (1993) compared three British surveys in terms of the technology applied. According to their results, CAPI takes more time in two surveys, whereas in the third study no significant difference was found. The increased length in CAPI was mainly attributed to the interviewers′ lack of experience.
- Finally, in a recent study Müller and Kesselmann (1996) compared PAPI, CAPI, CATI and CSAQ. They found no significant difference in terms of real interview duration (PAPI: 11.9 minutes; CAPI: 11.5 minutes); instead, there were interesting differences of the respondents perception of length of the interview (PAPI: 11.0 minutes, CAPI: 9.4 minutes).

By summarizing these results, we have no clear evidence that either technology is faster. In our view, these heterogeneous findings are due to the fact that the studies mentioned above do not contrast comparable tasks, that is, collect a specific amount of information with an identical instrument. In order to avoid this drawback we compare interview duration for single items that are comparable in terms of the amount of information collected. The main question of the present paper reads: which factors contribute to differences between CAPI and PAPI in terms of interview duration? Our focus in answering this question is on design issues and interviewer behaviors.

2 Methods

The comparison reported here is performed using the National Health Interview Survey (NHIS) instrument. The NHIS is conducted by the National Center for Health Statistic (NCHS) as an annual survey on the health conditions of the American noninstitutionalized population. It is an ongoing cross-sectional household interview survey of approximately 43,000 households including about 106,000 persons. The NHIS is conducted as a personal interview in the respondent′s home. It has completed the transition from PAPI to CAPI back in 1996/97.

In order to evaluate the computer assisted instrument used, the Survey Research Center in Ann Arbor, Michigan (USA) conducted a series of usability tests: in the spring of 1997, 37 CAPI interviews were performed by interviewers from the US

Bureau of the Census in the Survey Research Center's usability laboratory (for results related to usability issues see Hansen et al., 1997, 1998; Lepkowsky et al., 1998). We used the NHIS CAPI instrument programmed in CASES. It was installed on a standard laptop computer used by Census interviewers in the field. Moreover, 14 interviews applying the PAPI instrument used until 1996 were conducted in order to contrast the two technologies. 9 different interviewers from the Bureau of the Census' regional Office in Detroit were participating. Each interviewer took part in one or two sessions, each consisting of 6 to 8 interviews (4 to 6 CAPI and 2 PAPI). Respondents were recruited from the Ann Arbor area and paid for their participation. Minorities and women were represented. The maximum household size reached 6 persons leading to an average household size of 3.2. Respondents were randomly assigned to the CAPI or PAPI condition.

During the tests, three video tapes and one audio tape were recorded: one video tape showing the actual screen content of the laptop or the paper instrument, one recording the interviewer respondent interaction and a third one showing the interviewer's hands on the computer keyboard. In addition to those methods, we had respondent debriefings, trace files, and several other sources that documented the experimental cases. The main purpose of the analysis was developing and using a usability approach to evaluate the NHIS CAPI instrument.

- *searching* the next question on the screen or paper questionnaire (interviewer)
- *reading* the question text (interviewer)
- *negotiating* the meaning of question text and answer - asking for more information, providing an answer, probing, feedback, other task-related verbal and nonverbal contributions (interviewer and respondent)
- *recording* response (interviewer)
- *special* behaviors - digression, problems related to the questionnaire or the computer (interviewer and respondent)
- working with *tools* - calendars, booklets, etc. (interviewer)

Figure 1: Interviewer and respondent behaviors.

The results reported here are based on a separate coding path of the video tapes. More than 2,200 items from the socio-demographic portion of the instrument were coded. This segment of the instrument is very similar in the PAPI version as well as in the CAPI version (for details see below). Several relevant interviewer and respondent behaviors, as well as the duration, were coded for each item. Time was recorded based on the video tapes. For each behavior, the number of seconds elapsed was determined using the internal VCR clock.[4] For each item several time segments were

[4] Using this procedure we could avoid recording problems mentioned in another study (Lynn/Purdon, 1994: 152), namely the fact that the duration for CAPI interviews is recorded by the CAI system whereas under PAPI conditions interviewers are required to write down start time and ending time of the interview or interview section.

coded, each representing a specific interviewer behavior or respondent behavior of a certain length associated with the item. The following table lists the behaviors considered in the coding scheme. The data set contains a total of 7,031 segements. This approach allows a detailed assessment of the duration for every single item and helps detect differences in the participants' behaviors - even if the total duration of an item is equal under both conditions.

3 Results

3.1 Overall duration

Both instruments - the paper based as well as the computer assisted one - collect about the same amount of core information. Besides these identical items in both instruments, either questionnaire contains, however, some questions that were not implemented in the other version. For example, the paper instrument consists of check boxes and other items for controlling the flow of the interview, whereas the computer assisted instrument contains items for determining the respondent or for collecting information in greater detail than the paper instrument. For the purpose of the analysis reported in this paper, we focus on comparable items only. This means that the segments included in our data set collect the same amount of information in both instruments, and that the analysis does not include any other items.

Due to the fact that respondents were randomly assigned to the conditions, we could not make sure that we will have the same average household size for both conditions in the end. As a result of the random assignment, the average number of eligible persons living in a household reaches 2.9 for PAPI interviews and 3.3 for computer assisted interviews. Because of the NHIS collecting information for all eligible household members in its socio-demographic portion, we have to take into account these dissimilar values when computing the differences in terms of duration. Even thought the values do not differ statistically significant, the averages reported in Figure 2 are adjusted for household size.

On the average, the computer assisted interviews take almost 40 seconds longer than interviews conducted under the the paper based condition (230 compared to 268 seconds), which represents an increase of about 16.5 percent ($p < 0,05$). If we look at the two persons acting as participants in the interview situation - interviewer and respondent - we compute almost the same proportional differences for both partners. Interviewers, as well as respondents, spend more time during the course of the interview when using a computer assisted instrument. Due to the small sample size, the difference for respondents does not reach statistical significance. Still for the overall duration as well as for the time spent by interviewers, we do have differences that reach the level of significance.

	PAPI	CAPI	total
overall duration	230 sec	268 sec	258 sec *
interviewer time	163 sec	189 sec	182 sec *
respondent time	67 sec	80 sec	76 sec
question per household member	8.9	11.1	10.5 ***
duration per question	8.9	7.6	8.0 *
turns/actions per question	3.7	3.3	3.4 **

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; mean duration adjusted for household size.

Figure 2: Interview duration of the socio-economic portion and factors contribution to the interviewer duration (comparable items only).

If we assess the reasons for the longer duration of the computer assisted instrument, we face some interesting findings: in the paper based instrument, the interviewers ask on average 8.9 questions per household member to collect a specific amount of information, whereas they ask 11.1 questions under the CAPI condition to collect the same amount of information for every eligible person in the household ($p < 0.001$). On the other hand, the time spent on each of these items is significantly shorter in the computer assisted instrument. The interviewer spends 8.9 seconds in the paper based instrument compared to 7.6 seconds in the computer assisted questionnaire ($p < 0.05$). This can be explained in part by the fact that both participants in the interview situation take more action to complete their tasks in the paper-based instrument. Interviewer and respondent on the average need 3.7 turns or actions for every question in the paper-based instrument, but only 3.3 turns or actions in the computer assisted questionnaire ($p < 0.01$).

Looking at the means for the overall duration only does ignore the differences taking place on a more detailed level. Some of these differences increase the duration for the computer assisted instrument, while other effects contribute to a longer duration of the paper based questionnaire. The remainder of the present paper deals with the contribution of these technology effects to the length of both versions.

3.2 Loop design instead of filters

One of the main differences in terms of questionnaire design relates to the use of filter questions. Whereas the paper based instrument usually applies a filter structure and several follow-up questions - if appropriate -, the computer assisted instrument generally uses a loop of questions for every eligible person in the household to obtain the information. This difference shall be explained by means of using the armed

forces question in both versions. The goal of this segment is to determine which of the adult household members is eligible as respondent or sample adult in the course of the interview. Persons serving in the armed forces are not eligible. In the paper instrument, a filter question works as a main gate for finding out whether or not any person in the household is on full time active duty with the armed forces. If the respondent provides the information that any of the persons is currently serving, a series of follow-up questions is asked, which person it is and whether or not this characteristic applies to any other adults in the household. The computer assisted instrument on the other hand uses a separate item asking for every single person in the household, whether or not they are on full time active duty with the armed forces (Figure 3).

This different question logic - that applies to several sections in the questionnaire- is the main reason for the larger number of items within the computer assisted instrument. The system needs one question for the whole household (plus follow-up) in order to collect the appropriate information in the paper based instrument. In contrast, we need one question for every person in the household in a CAPI instrument (plus follow-up), which amounts to more than two questions for the average household with 2.3 eligible persons (the question applies to adults of a certain range of age, only).

However, a questionnaire applying a filter design needs a more thorough administration by the interviewer and implies an extended cognitive burden for the respondent. When answering the filter question in the PAPI instrument, the respondent has to take into account all adults in the household at once. On the other hand, by answering the single items in the loop design, the respondents′ focus is already set on the person in question, and they do not have to worry about any other person in the household at this time. Looking at the duration of the armed forces question (Figure 4), we find supporting evidence for our interpretation. The overall duration of the filter design is substantially shorter (2.4 seconds) than the combined duration of all single questions in the CAPI instrument applying a loop design. In comparison, answering every single question in the loop design takes much less time than answering the filter question in the PAPI version. Consequently, for small households with just one eligible adult to whom the armed forces question applies: the loop design is still faster than the filter design implemented in the paper based instrument. For larger households with three to five household members, the sum of time for every single item leads to a total duration covering a larger period than in the filter design ($p < 0.001$).

From these findings we can draw the conclusion that the average number of entities to whom a specific question applies is one important factor for determining whether or not a loop design is more appropriate than a filter question. If there is only one eligible person in the household, the CAPI version runs faster. Another key factor is the expected proportion of those entities, who do fulfill the characteristics in question. Assuming that only a very small proportion of all eligible household members serves on full time active duty with the armed forces, it definitely makes sense to apply a filter design because the number of instances where the interviewer has to

work through the follow up sequence is relatively small. In addition, under these conditions the loop design wastes a lot of time for questions that do not lead to a high variation in the data set.

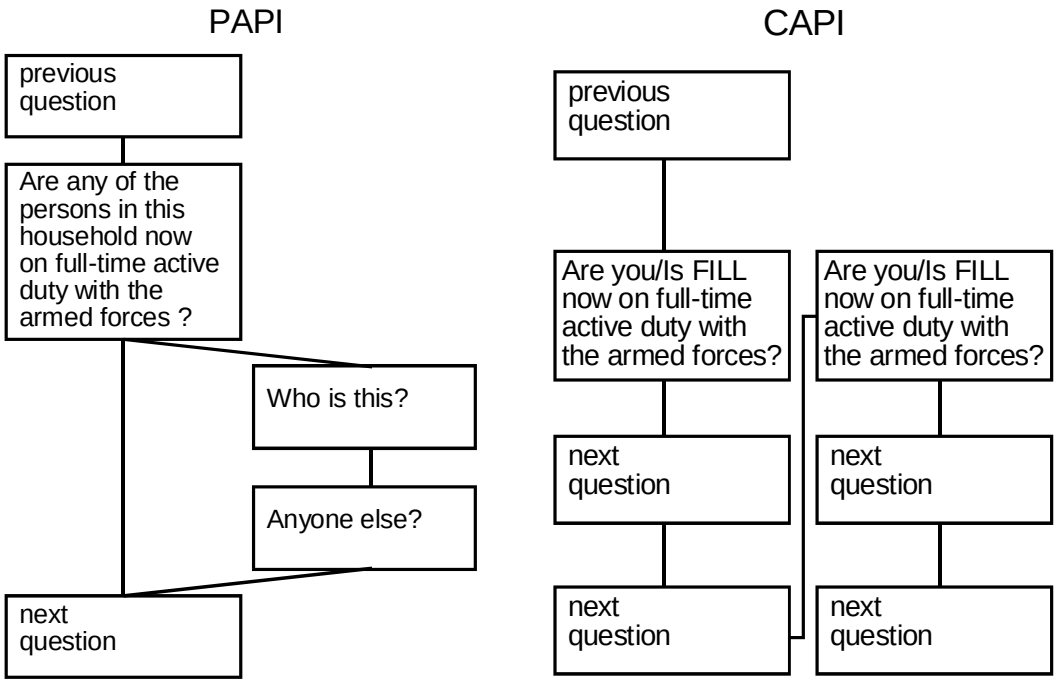

Figure 3: Loop design and filter logic applied to the armed forces question.

Applying a filter question or a loop design to an instrument is not a technology effect in the first place. However, by looking at some instruments, we find that paper questionnaires usually apply a filter design in order to determine whether or not a specific entity within a household fulfills a certain characteristic. Computer assisted instruments, by contrast, usually make use of a loop design. At least two reasons connected to computer technology can account for this phenomenon:

1. On the one hand, a questionnaire designed as a computer assisted instrument - at least in part - switches control over the flow of the interview from the interviewer onto the computer. The computer applies a very rigid question order and allows only little flexibility in the actual interview situation in terms of adapting to its specifics. Thus, the designer of a computer assisted instrument cannot anticipate every possible variation of the interview flow, so that he or she has to make sure that the instrument covers all possible circumstances in a person′s household. To ascertain that the computer assisted instrument collects high quality data, its design has to take into account the necessity that in-

formation be assembled on a highly disaggregated level: one bit of information for every single person at a time.

2. On the other hand, a loop design matches the way in which computer assisted instruments treat entities within a household. Most CATI or CAPI systems understand persons in a household as a subrecord of the compounded observation called household. In the background of computer assisted instruments works a data base that treats households as the main entry and entities like person as a dependent subentries. From a programmer's point of view, it is a natural and easy way to accumulate information for these entities in a loop design.

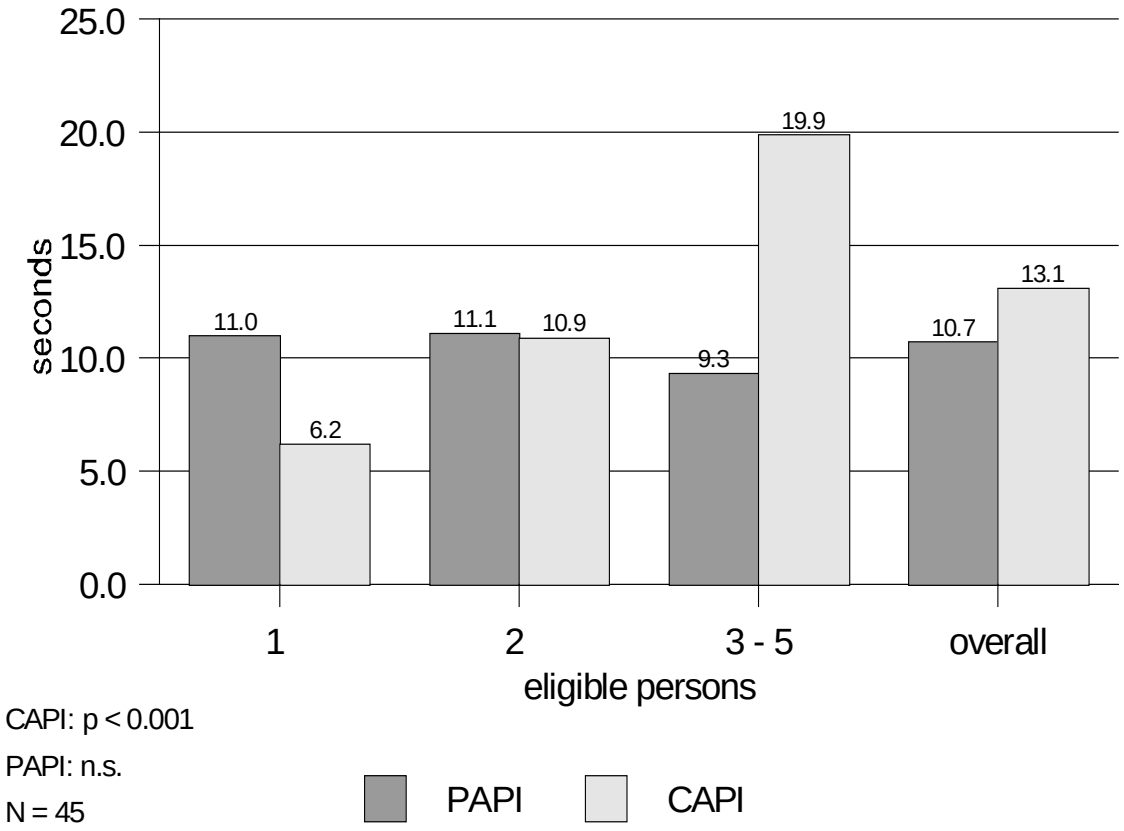

Figure 4: Duration of "armed forces" question (per interview) by technology and household size.

3.3 Character input and banked screens

Another difference between PAPI and CAPI directly connected to the technology applied, results from the use of a keyboard instead of a pen for recording the information provided by the respondent. Looking at the very first substantial question of the instrument, we can demonstrate the two effects related to this change of technology. The questions reads as: "What are the names of all persons living or staying here? Start with the name of the person or one of these persons who owns or rents this home." The question wording is identical in both versions and even the amount of information required is the same. The respondent is expected to reveal the first name, middle initial, and last names of all persons living in the household. Figure 5 shows a screen dump from the CAPI version.

The results of the analysis show that it takes substantially more time ($p < 0.001$) to complete this question in the computer assisted version compared to the paper instrument (CAPI: 20.6 seconds, PAPI: 17.5 seconds). The values reported in the Figure 6 represent the average duration of several interviewer and respondent behaviors for each instance this question is asked. The results show no significant difference in terms of time for the interviewer searching the next question and preparing for delivering it. Also, interviewers spend the same amount of time for reading the question text under both conditions. Moreover, there is no significant difference between the two versions in terms of time spent on additional actions or behaviors like digression or dealing with problems related to the questionnaire or the computer and so on ("special").

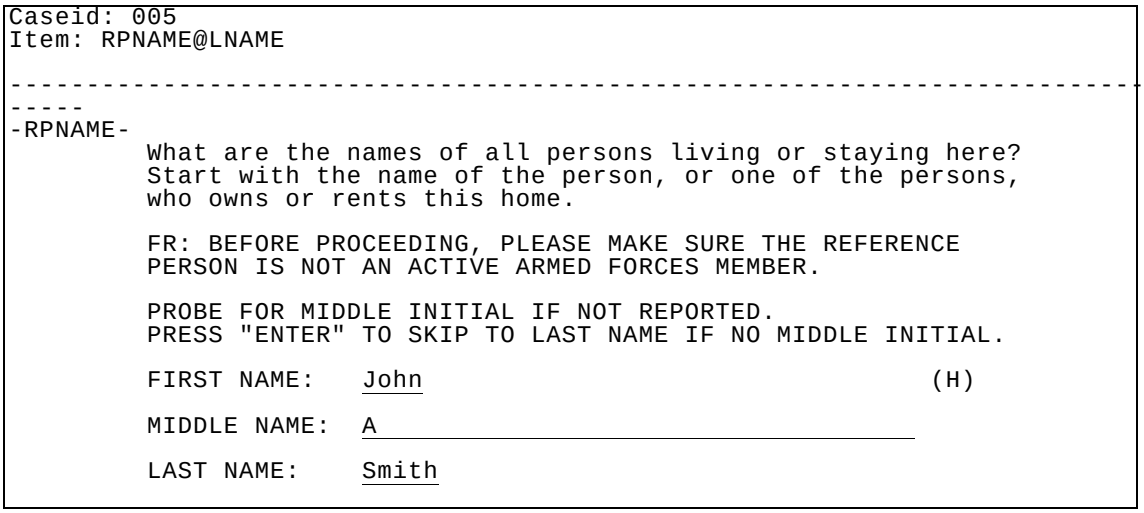

Figure 5: Example for string input on a multiple item screen (NHIS instrument, CAI version).

It is interesting to observe that the negotiations between interviewer and respondent (asking for more information, providing an answer, probing, providing feedback, other task related verbal contributions and non-verbal turns) take substantially longer

in the computer assisted version as compared to the PAPI version. This is mainly due to the different degree of flexibility provided by either version. The computer assisted instrument is designed as a multiple item screen combining several input fields on one screen. There are three input fields on this screen: first name, middle initial, and last name. The system expects the input in a certain order: first name, middle initial, and last name. In case the respondent provides the last name first, the interviewers have two options for dealing with this situation: (1) they can type dummy information for the first name and press [enter] for the middle initial in order to move the cursor to the third input field. Then, they can record the last name and return to the first input field (2 x [backup]) to type the first name and the middle initial. (2) As a second option they can ask the respondent to provide the information in the order expected by the computer system. Whatever solution they choose, it takes more time compared to the paper instrument, where the interviewer points the pen to the appropriate blank space on the questionnaire form regardless of the order in which the respondent provides the answers.

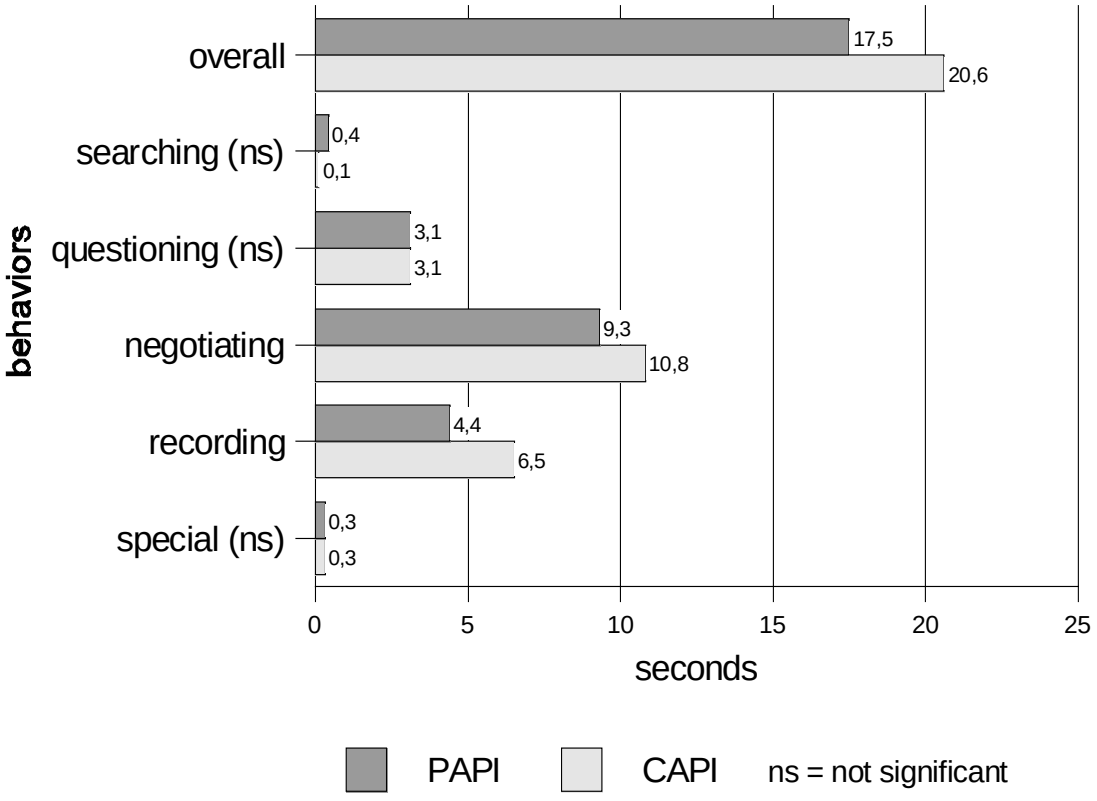

Figure 6: Duration of "What are the names …?" segment.

The second difference between the two questionnaire versions in terms of time is related to the input device, too. On the average it takes 4.4 seconds to write down the first name, middle initial, and last name of a person with a pen, whereas it takes 6.5 seconds to record this information using the keyboard. Thus, we can prove the

hypothesis that writing is faster than typing, even though, interviewers are experienced in typing names. This could be explained in part by the fact that, when writing, interviewers make fewer or no errors, but it is the correction of errors when typing which takes longer. This hypothesis is difficult to prove: we do know the number of corrections and back-ups in the CAPI version. However, we did not code these instances for PAPI interviews. Thus, based on the available data we cannot decide which explanation is true.

Furthermore, the difference in terms of time for recording may be explained by the fact that interviewers are allowed to edit the paper questionnaire after the completion of the whole instrument based on a handwritten draft used during the course of the interview. Instead, a completed computer assisted interview is no longer accessible to the interviewers once they have stored the information and closed the case. It is hard to think of any solution for this disadvantage of CAPI instruments. In fact, it does not make sense to type a name or other kind of string information in a less accurate or fast manner and allow interviewers to edit the input after the completion of the interviewer. However, providing the interviewer with advanced editing functions during the course of the interview and more flexible cursor movement on screens with multiple input fields might help speed up recording of character information.

3.4 Automated calculations and fills

The technology effect discussed in this section shows one of the advantages of computer assisted interviewing, namely the feature of making use of automated calculations and fills. In the socio-demographic portion of our instrument, the interviewer is required to collect the date of birth for all persons living in the household. In the paper questionnaire, he or she is forced to use a paper form (similar to a calendar) to calculate the age from the given date of birth and the actual interview date. By comparison the same task is less complicated in the computer assisted instrument: the system calculates the correct age and automatically builds a confirmation text that includes the appropriate name fill and the result of the computation: "That would make John Smith 33 years old. Is that correct? yes, no".

The results show (Figure 6) that this solution takes less than half of the time necessary to complete this task in the paper instrument (4.6 vs. 10.5 seconds). This difference is mainly due to the fact that the interviewer spends about four seconds on working with the tool to compute the age from the given interview date and the respondent's date of birth and also to the fact that he or she needs to write down a two-digit number instead of pressing a single key for yes or no in the computer assisted instrument. Additionally, there is a greater amount of the negotiation related to that task. As a result, the PAPI instrument requires more interviewer action and interaction between the interviewer and the respondent. Even though this might be time effective, we have to consider that the CAPI instrument requires this procedure

for every single household member, even if the appropriate information is provided by the respondent right away, along with the date of birth. As a result, the PAPI version might be even shorter under these conditions.

On the other hand, not applying automated calculations and fills to the CAPI instrument would increase the duration even more. When designing a computer assisted instrument, we should consider other questions or items where we could make use of this advantage extensively. Besides this advantage in terms of duration, the computer assisted solution helps improve data quality, too: it confirms the calculated age and produces an additional degree of certainty that the recorded information is correct.

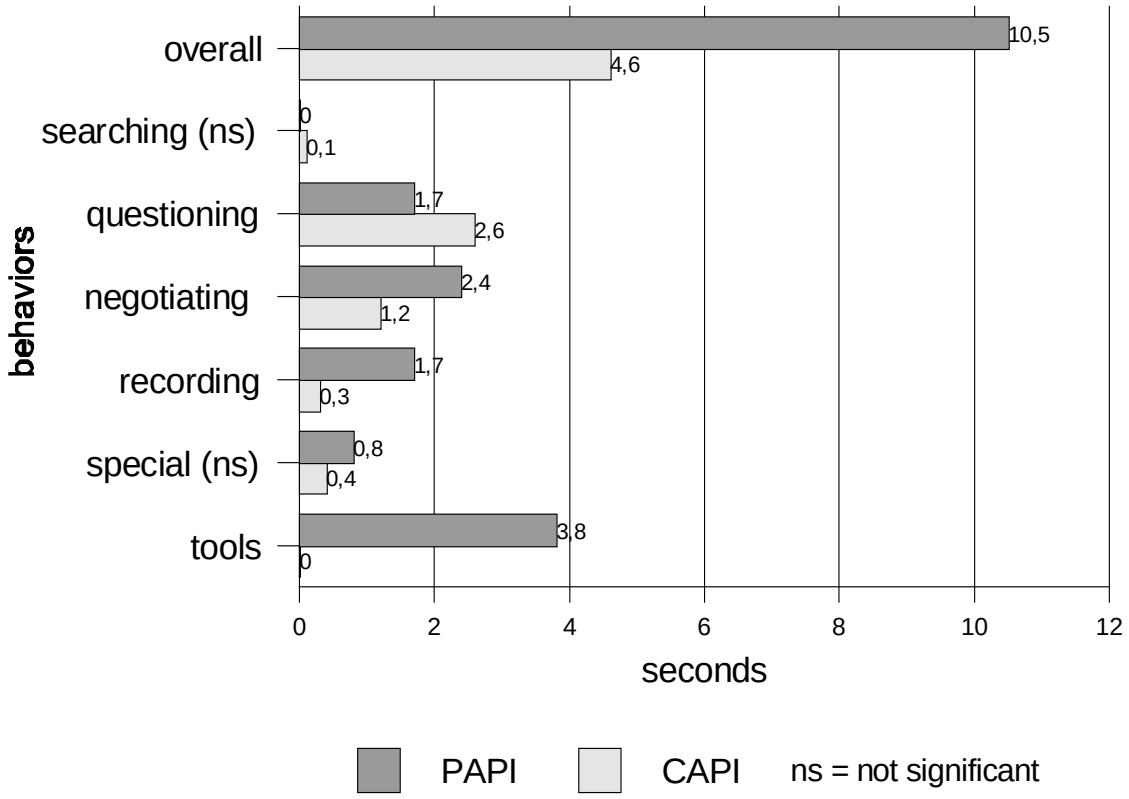

Figure 7: Duration of "age" item.

We have shown before that recording alpha input using a keyboard takes more time than working with a pencil. We could hardly think of any solution for that disadvantage of CAI. Even though we cannot avoid this drawback of technology, we can try to cover this increase of time by implementing features that help spare time. Making use of automated calculations and fills is one way to reduce the amount of time necessary to complete a task. Using questions like the age item discussed in this section even has an additional positive influence on the interview situation. Making use of the computer's ability to store and compute information, we can develop more intelligent instruments that demonstrate the CAI system's customized appearance and support the interviewer's responsiveness.

3.5 ”Real” comparison

So far we have discussed several differences between the PAPI and the CAPI version, that are directly or indirectly connected to the technology applied: a loop design instead of filters, making use of a keyboard instead of a pencil, and using the computer′s ability to perform calculations and its capabilities to make use of information previously collected. In this section we want to compare a series of questions that is identically worded in both instruments and that applies the same question logic and expects the same responses or inputs from the interviewer. None of the technology effects previously discussed characterizes this segment.

To assess any additional differences we compare the household composition probe: After the completion of the household roster, the interviewer is required to make sure that he or she has covered the whole household and does not ignore or forget any additional persons. In order to verify the household composition he or she probes: ″I have listed as living here ...″, then he or she reads the names of all persons living in the household, and then he or she asks ″... have I missed any babies or small children? Any lodgers, boarders, or person you employ who live here? Anyone who usually lives here but is now away from home, travelling or in a hospital? Anyone else staying here?″ The questionnaire expects a simple yes or no answer for each item.

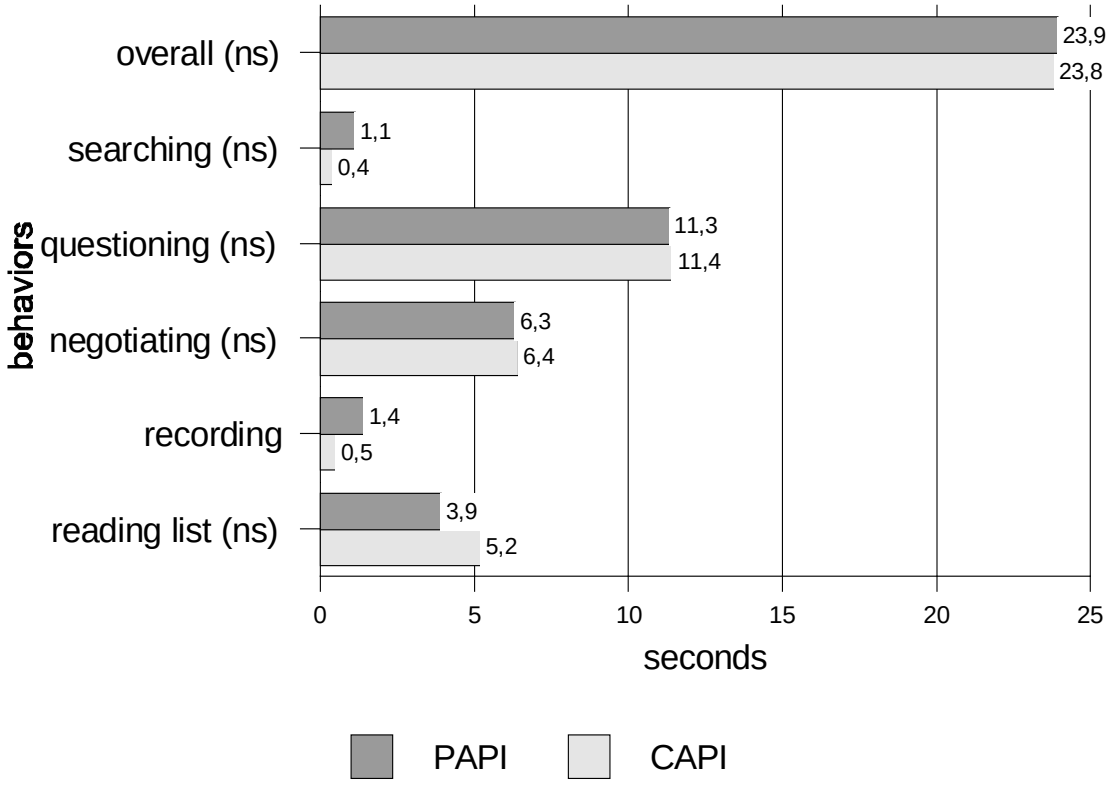

Figure 8: Duration of “missing person” segment.

Looking at the results (Figure 8), we learn that it makes no difference whether or not this household composition probe is administered using a computer or a paper instrument. The overall duration is identical and we do not observe significant differences for searching the next question, delivering the question to the respondent, negotiating the meaning of the question or the respondent's answer, and reading the list of all persons living in that household. The only difference that reaches a level of statistical significance relates to the process of recording the appropriate information: it takes more time to mark the four check boxes on a paper form using a pen than it takes to put down four single digits (1 = yes, 2 = no) and pressing [enter] in the computer assisted instrument. Apart from these findings, we have learned that the computer does not make any difference in terms of time if the compared sequence is not affected by other features introduced by the computer technology.

4 Discussion

The data reported in this paper support the hypothesis that the introduction of the laptop computer into the interview situation does not contribute to a longer duration of the interview as such. Still, the technology involved in a CAPI instrument comes into effect. (1) The overall interview duration can benefit from the system's possibility to compute calculations automatically and make use of fills and other information previously recorded or drawn from external data files. (2) Typing alpha-numeric information using a keyboard takes more time than writing down comparable information with a pencil. (3) Finally, we could show that the loop design usually applied in CAPI instruments consumes more time than the filter logic that is used in many of the paper based questionnaires. As a result of these factors contributing to the interview duration, the CAPI instrument takes substantially more time to collect the same amount of information (+16.5 %). In contrast, we could demonstrate that it does not make any difference in terms of duration if we compare a sequence of questions that is identical by any means in both versions.

It is not yet clear whether or not a longer duration of an interview is a problem by itself. It might be even favorable if we achieve better data quality due to a fool-proof design, a rigid question order, and other features. However, a longer interview duration implies the risk of a higher break-off rate. It is well known that refusals occur mainly at the beginning of an interview, but we also know that the interview duration and the respondent's satisfaction, as well as the attitude toward the interviewer contribute to the break-off rate, too. Finally, we have to consider that the additional interview duration contributes to the total survey cost.

Additionally, there is another drawback connected to the technology: lack of flexibility. Our results from another experiment (not reported in this paper; see Fuchs, 1998) show that the fool-proof rigid question order implies a certain degree of non-standard interview behavior. Interviewers try to make the questionnaire work in the

interview situation and use some deviations from standard interviewer techniques to administer questions that would be perceived as redundant or embarrassing by the respondent. For example, the respondent might provide some relevant information along with the answer to another question. When using a CAPI instrument, the interviewer has no chance to record it before the appropriate screen comes up. If he or she comes to the point where the information is to be stored, the interviewer has two options. (1) He or she can read the question as worded. That would be perceived as unresponsive and redundant by the respondent. (2) He or she can customize the question according to the degree of information previously provided by the respondent. The interviewer can confirm the information or record the answer without reading the question at all. As a result, the rigid question order leads to more deviations from standard interviewer procedures or to an interview that appears to be a less customized and responsive. From the researcher's point of view neither effect is favorable.

Therefore, our results contribute to a more general discussion about the usability of computer assisted instruments, too. We have tried to show that technology influences the interview flow, the question logic, and the screen design as well as the interviewer's interaction with the instrument. Also, we have tried to demonstrate that some CAI design solutions and screen layouts do not support interviewers in efficiently solving their tasks in the interview situation. This leads to a more general requirement: a CAPI instrument needs a careful assessment of its usability to make sure that we do benefit from its advantages in terms of data quality without suffering from its disadvantages caused by poorly designed screens and features. Not everything that can be programmed and appears nice on the screen works well in the field.

References

[1] Baker, R.P. (1992): New Technology in Survey Research: Computer-Assisted Personal Interviewing (CAPI). *Social Science Computer Review*, **10**, 145-157.

[2] Baker, R.P., Bradburn, N., and Johnson, R. (1994): CAPI: An Experimental Evaluation. In American Statistical Association (Ed.), *Proceedings of the Section on Survey Research Methods*, 851-855.

[3] Booth-Kewley, S., Rosenfeld, P., and Edwards, J.E. (1993): Computer-Administered Surveys in Organizational Settings. Alternative, Advantages and Applications. In Rosenfeld, P. et al. (Eds.), *Improving Organizational Surveys*, Newbury, 73-101.

[4] 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 Groves, R.M., Biemer, P. P., Lyberg, L. E., Massey, J. T., Nicholls II, W. L., and Waksberg, J. (Eds.), *Telephone Survey Methodology*, New York, 437-450.

[5] Couper, M. P., Groves, R. M., and Kosary, C. (1989): Methodological Issues in CAPI. In American Statistical Association (Ed.), *Proceedings of the Section on Survey Research,* 349-354.

[6] Couper, M.P. and Groves, R.M. (1992): Interviewer reactions to alternative hardware for computer-assisted personal interviewing. *Journal of Official Statistics*, **8**, 201-210.

[7] Couper, M.P. (1994): What can CAI learn from HCI? Paper presented at the COPAFS seminar, June 1994.

[8] Couper, M.P. and Burt, G. (1994): Interviewer Attitudes toward Computer-Assisted Personal Interviewing (CAPI). *Social Science Computer Review*, **12**, 38-54.

[9] Couper, M.P. (1997): The Application of Cognitive Science to Computer Assisted Interviewing. Paper Presented at the CASM II Seminar, June 12, 1997.

[10] Couper, M.P., Fuchs, M., Hansen, S.E., and Sparks, P. (1997a): CAPI Instrument Design for the Consumer Expenditure (CE*) Quarterly Interview Survey. Final Report*. University of Michigan.

[11] Fuchs, M. (1994): Umfrageforschung mit Telefon und Computer. Einführung in die computergestützte telefonische Befragung. Weinheim: Psychologie Verlags Union.

[12] Fuchs, M. (1995): Die computergestützte telefonische Befragung. Einige Antworten auf Probleme der Umfrageforschung. *Zeitschrift für Soziologie*, **24**, 284-299.

[13] Fuchs, M. (1997): Interviewer behavior and Design Issues in a CATI Survey. Keystroke Files from the Detroit Area Study. Paper presented at the Brown Bag Seminar, Survey Reseach Center, Ann Arbor, USA, May 1997.

[14] Fuchs, M. (1998): CAI Screen Design and Interviewer Behavior-Results from a CATI Field Experiment. Paper Presented at the SRC's Brown Bag Seminar, Ann Arbor, USA, May 1998.

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

[16] Hansen, S.E., Fuchs, M., and Couper, M. (1997): CAI Instrument Usability Testing. Paper Presented at the Annual Meeting of the American Association for Public Opinion Research, Norfolk, May 1997.

[17] Hansen, S.E., Couper, M.P., and Fuchs M. (1998): Usability Evaluation of the NHIS Instrument. Paper presented at the Annual Meetings of the AAPOR, St. Louis, MO, May 1998.

[18] Harlow, B.L., Rosenthal, J.F., and Ziegler, R.G. (1985): A Comparison of Computer-Assisted and Hard Copy Telephone Interviewing. *American Journal of Epidemiology*, **122**, 335-340.