Journal of Official Statistics, Vol. 16, No. 3, 2000, pp. 273–286

Technology Effects: Do CAPI or PAPI Interviews Take Longer?

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

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 article 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 to the technology itself.

Since the early days of computer-assisted interviewing (CAI), the issue of whether computer-assisted interviews take longer than equivalent paper and pencil surveys has been discussed. Evidence in both directions can be found in the literature. To what extent does the time taken to administer survey questions depend on the technology (paper or computer) used? Or does the duration of an interview depend on the particular design features employed? In this article we explore this issue in detail, at the level of individual questions. Our goal is to understand whether and if so why the time taken to complete items on paper and on computer may differ.

Key words: CAPI; interview length; instrument design; computer-assisted interviewing.

1. Interview Duration in a CAI Environment

What do we know about CAI interview duration? Early studies conducted in the 1980s found that computer-assisted telephone interviewing (CATI) took longer than paper-based

[1] Catholic University of Eichstaett, Department of Sociology, 85071 Eichstätt, Germany. E-mail: marek.fuchs@ku-eichstaett.de

[2] University of Michigan, Institute for Social Research, 426 Thompson St., Ann Arbor, MI 48104, U.S.A. E-mails: MCouper@umich.edu and SEHansen@umich.edu

Acknowledgments: The work reported here was financed by a research stipend from the Humboldt Foundation of Germany to the first author, who was a visiting scholar at the Survey Research Center, University of Michigan during 1996/97. The usability tests on which this analysis is based were conducted under a cooperative agreement (principal investigator: James M. Lepkowski) supported by the U.S. National Center for Health Statistics. The authors would like to thank the associate editor William L. Nicholls II and several anonymous reviewers for their helpful suggestions regarding an earlier draft of this article.

© Statistics Sweden

interviewing (Catlin and Ingram 1988; Groves and Mathiowetz 1984; Harlow, Rosenthal, and Ziegler 1985; see also Nicholls and Groves 1986). The time differences were mainly attributed to hardware and software limitations at the time. The results for computer-assisted personal interviewing (CAPI) are somewhat more mixed. An early study by Birkett (1988) found that CAPI interviews took about 25% longer than the equivalent paper interviews (see Waterton and Duffy 1984, for related results). Baker, Bradburn, and Johnson (1995) found that CAPI interviews in the National Longitudinal Survey/ Youth Cohort (NLS/Y) lasted an average of 47 minutes, compared to 57 minutes for paper. They also noted that the completion time of CAPI interviews declined steeply over the field period (from 52 to 42 minutes), while the decline in the length of the paper interviews was not as steep. However, the time measures used in their survey did not include the household enumeration section. Furthermore, the NLS/Y is a very complex instrument, with over 60 interviewer instructions to transcribe information from the cover sheet into the interview (all done automatically in CAPI) and a number of arithmetic calculations (again automated in CAPI). Martin, O'Muircheartaigh, and Curtice (1993) compared three British surveys conducted using paper and pencil interviewing (PAPI) and CAPI. CAPI took significantly longer in the first two surveys, but the difference was not significant in the third study. They attributed this to increasing interviewer experience with CAPI. This conclusion is supported by Lynn (1998; see also Lynn and Purdon 1994), who reported an interview length of 45.5 minutes for CAPI in the British Social Attitudes Survey, compared to 54.0 minutes for PAPI. Müller and Kesselmann (1996) compared PAPI, CAPI, CATI and computerized self-administered questionnaires (CSAQ). They found no significant differences between PAPI and CAPI in terms of interview duration (11.9 minutes for PAPI and 11.5 for CAPI). Interestingly, respondents perceived the PAPI interview as taking longer (an average estimated time of 11.0 minutes for PAPI and 9.4 for CAPI).

What may account for these different findings on relative length of CAI and paper interviews? Certainly, the speed of the computer hardware may have been a factor in the early days of CAI, but with the modern laptops and desktops used for today's interviewing, this is unlikely to be a major factor. Secondly, the design or implementation of CAI software may limit the speed of the interview. Many CAI surveys make use of a single question per screen approach, especially when they are DOS-based systems (this may change with the advent of Windows-based interviewing systems). This design feature may slow the interview down relative to a paper and pencil interview in which the interviewer can see many questions at once on a page. Thirdly, the speed of writing versus typing may favor PAPI over CAPI, but only for those surveys which require a great deal of text entry. Finally, possible difficulties reading a laptop computer screen in less-than-ideal conditions may have a detrimental effect on the speed of CAPI interviews.

On the other hand, there are several factors favoring CAI in terms of interview speed. Automated skips or branching reduce the time taken by interviewers to figure out what question to ask next. Automated arithmetic calculations are likely to be much faster than when done by the interviewer. Customized text delivery using fills may reduce interviewer time and effort to choose the appropriate wording. Automated consistency and edit checks may similarly reduce interviewer time doing such checking, if equivalent checks have been built into the paper instrument. On the other hand, if such checks trigger a consistency resolution, the interview may be slowed down.

In summary, then, we expect that when the instrument is complex, with many contingent questions, branching decisions, calculations, and few open-ended items requiring text entry, the time comparison will favor CAI. On the other hand, an instrument with few edits or branching decisions, and with many similar questions (e.g., a series of attitude items) presented on separate screens in CAI, the time comparison may favor paper. In other words, it is the design of the survey and instrument rather than the technology that determines the time taken to complete an interview. In this article we examine parts of an instrument that contains many of the features outlined above. This permits detailed comparison of the possible sources of time differences between paper-based and computer-assisted interviews.

2. Methods

The comparison of question durations between PAPI and CAPI is based on the National Health Interview Survey (NHIS) instrument. The NHIS is conducted by the U.S. Bureau of the Census on behalf of the National Center for Health Statistic (NCHS), and focuses on the health conditions of the household population of the United States. It is an ongoing cross-sectional household interview survey of approximately 43,000 households containing about 106,000 persons. The NHIS is conducted as personal interviews in the respondents' homes. CAPI was introduced in parallel with the paper survey during 1996, and by January 1997 the conversion of the NHIS to CAPI was complete.

In order to evaluate the computer-assisted instrument used, the Survey Research Center (SRC) at the University of Michigan conducted a series of usability tests in the Spring of 1997. A total of 37 CAPI interviews were conducted by interviewers from the U.S. Bureau of the Census in SRC's usability laboratory (see Hansen, Fuchs, and Couper 1997; Hansen, Couper, and Fuchs 1998; Lepkowski et al. 1998). The NHIS CAPI instrument was programmed using CASES 4.3. It was installed on a standard laptop computer used by Census interviewers in the field. The same interviewers also conducted 14 interviews using the PAPI instrument. Nine interviewers participated in the test, with each interviewer doing 2 to 6 CAPI and 1 or 2 PAPI interviews. All the interviewers had experience on both the paper and pencil and CAPI instruments, having used the latter for at least six months prior to the usability tests. Respondents were recruited from the area and paid for their participation. The interviewers conducted their second interview each day on PAPI, so there was no self-selection in the assignment of interviewers or respondents to technology.

Several video tapes were recorded for each interview (see Hansen, Fuchs, and Couper 1997). The results reported here are based on a separate coding of the video tapes. More than 2,200 items from the household listing portion of the instrument (containing sociodemographic measures) were coded. This segment of the instrument is very similar across the two versions, permitting comparison of question duration. Several relevant interviewer and respondent behaviors, as well as the time 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. The same timing approach was used for both PAPI and CAPI interviews, permitting direct comparisons across modes. Figure 1 lists the behaviors considered in the coding scheme. The data set contains a total of 7,031 segments (subsets of items). This approach allows a detailed assessment

- *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 non-verbal 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)

Fig. 1. Interviewer and respondent behaviors

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 in both technology conditions.

3. Results

3.1. Overall duration

Both the paper-based and computer-assisted instruments collect about the same amount of core information. In addition to the common items, each instrument also contains items not implemented in the other version. For example, the paper instrument has check boxes and other items for controlling the flow of the interview, whereas the CAPI instrument contains items for determining the selected respondent or for collecting information in greater detail than the paper instrument. For the purpose of the analysis reported in this article, we focus on comparable items only. The segments included in our data set thus collect the same amount and type of information in both instruments.

We did not control for differences in household size in the assignment of respondents to the conditions. As a result, the average household size was 2.9 for the PAPI interviews and 3.3 for CAPI. Because the NHIS collects socio-demographic information on all eligible household members, we need to take these differences into account when examining time differences. Consequently, the averages reported in Figure 2 are adjusted for household size.

On average, the CAPI interviews take 38 seconds longer ($p < 0.05$) than interviews conducted under the paper-based condition, which represents an increase of about 16.5% over paper. If we look at the time taken by the two persons acting as participants in the interview situation – interviewer and respondent – we compute almost the same proportional differences for each; both interviewers and respondents spend more time

	PAPI	CAPI	Difference
Overall duration	230 sec	268 sec	38 sec*
Interviewer time	163 sec	189 sec	24 sec*
Respondent time	67 sec	80 sec	13 sec
Questions per household member	8.9	11.1	2.2***
Duration per question	8.9 sec	7.6 sec	−1.3 sec*
Turns/actions per question	3.7	3.3	−0.4**

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

Fig. 2. Interview duration of the socio-demographic section and factors contributing to the interview duration (comparable items only)

when using a computer-assisted instrument (given the small sample size, the difference for respondent time does not reach statistical significance).

If we assess the reasons for the longer duration of the computer-assisted instrument, we find some interesting differences. In the paper instrument, interviewers ask an average of 8.9 questions per household member for this series, compared to 11.1 in CAPI to collect the same information ($p < 0.001$). On the other hand, the time spent on each of these items is significantly shorter in the computer-assisted instrument. It takes an average of 8.9 seconds per question in the PAPI instrument compared to 7.6 seconds per item in the CAPI instrument ($p < 0.05$). This is explained in part by the fact that the two interactants take more conversational turns or actions to complete the task in the paper version (3.7) than in CAPI (3.3 turns or actions).

Simply comparing the overall duration for a set of items ignores the differences taking place on a more detailed level. The remainder of this paper deals with the contribution of technology and design effects to the relative length of survey instruments.

3.2. Loop design versus filter questions

One of the main differences between the PAPI and CAPI versions of the NHIS relates to the use of filter questions. Whereas the paper-based instrument usually applies a filter structure with several follow-up questions (where appropriate), the CAPI instrument generally uses a loop of questions for every eligible person in the household. This difference is illustrated using the armed forces question in both versions. The goal of this segment is to determine which household members are eligible for the remainder of the interview. Persons serving in the armed forces are not eligible. In the paper instrument, a single filter question asks whether any person in the household is on full-time active duty with the armed forces. If any household member is currently serving, a series of follow-up questions is asked. The computer-assisted instrument asks this question of every single adult in the household. The difference in question structure is illustrated in Figure 3. This difference in question logic (which applies to several sections of the NHIS questionnaire) is the main reason for the larger number of questions per household member in the CAPI instrument reported in Figure 2.

However, a questionnaire applying a filter design may require additional effort on the part of the interviewer to process the household-level wording and apply it appropriately for a single-person household. This may account for the longer time for the filter question in single-person households (see Figure 4). On the other hand, for larger households (3–5 members) the total time for these questions in a loop is longer than the time taken using the filter design ($p < 0.001$). Overall, the filter design of the paper instrument takes less time (2.4 seconds) than the combined duration of the set of person-level questions in the CAPI instrument applying a loop design.

From these findings we can conclude that the average number of entities (persons, in this case) to which a specific question applies is one important factor in 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 is faster. Another key factor is the expected proportion of entities which fulfill the characteristics in question. Assuming that only a very small proportion of all eligible household members are on active duty, it

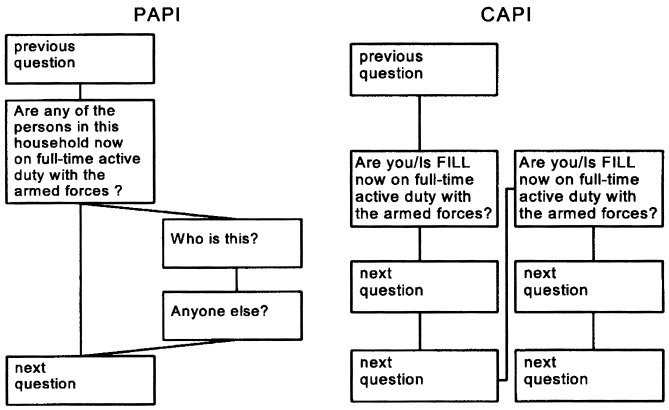

Fig. 3. Loop design and filter logic applied to the "armed forces" question

makes more sense to apply a filter design because the number of instances where the follow-up sequence is triggered is relatively small. Under these conditions the loop design wastes time asking questions that generally do not yield a positive response.

Applying a filter question or a loop design to an instrument is not necessarily an effect of technology (paper versus computer). In fact, some questions in the CAPI version of the NHIS employ a filter-based approach. However, loop designs appear more common in CAI surveys. At least two reasons connected to computer technology can account for this phenomenon:

(1) A computer-assisted instrument typically switches control over the flow of the interview from the interviewer to the computer. The computer applies a rigid question order and allows the interviewer little flexibility in the conduct of the interview. The designer of a CAI instrument cannot anticipate every possible variation of the interview flow, so needs to ensure that all possible eventualities are covered. This leads to highly disaggregated data collection: one item of information for each person, one at a time.

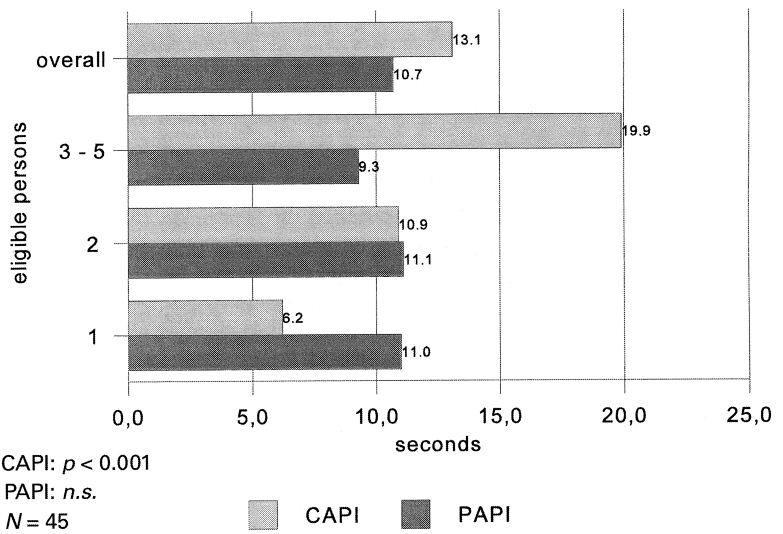

Fig. 4. Duration of "armed forces" questions (per interview) by technology and household size

(2) A loop design matches the way in which computer-assisted instruments treat entities within a household. Most CAI systems treat persons in a household as a subrecord of the household in a hierarchical database. From a programming perspective, it is natural and easy to accumulate information for these entities in a loop design. In other words, the CAI instrument anticipates the structure of the database produced, which is not necessarily the case on paper.

3.3. Text input and multiple item screens

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

We find that it takes substantially more time ($p < 0.001$) to complete this question in the computer-assisted version (20.6 seconds) compared to the paper instrument (17.5 seconds). The values reported in Figure 6 represent the average duration of several interviewer and respondent behaviors for each instance where this question is asked, as well as the overall time. The results show no significant difference in terms of time for the interviewer to find the next question and prepare to deliver it. Also, interviewers spend the same amount of time reading the question text under both conditions (3.1 seconds). Moreover, there is no significant difference between the two versions in terms of time spent on ''special'' behaviors like digressions or dealing with problems related to the questionnaire or the computer, and so on.

```
Caseid: 005
Item: RPNAME@LNAME

--------------------------------------------------------------------------------
-RPNAME-
         What are the names of all persons living or staying here?
         Start with the name of the person, or one of the persons,
         who owns or rents this home.

         FR: BEFORE PROCEEDING, PLEASE MAKE SURE THE REFERENCE
         PERSON IS NOT AN ACTIVE ARMED FORCES MEMBER.

         PROBE FOR MIDDLE INITIAL IF NOT REPORTED.
         PRESS "ENTER" TO SKIP TO LAST NAME IF NO MIDDLE INITIAL.

         FIRST NAME:    John                                    (H)

         MIDDLE NAME:   A

         LAST NAME:     Smith
```

Fig. 5. Example of text input on a multiple item screen (CAPI version of NHIS)

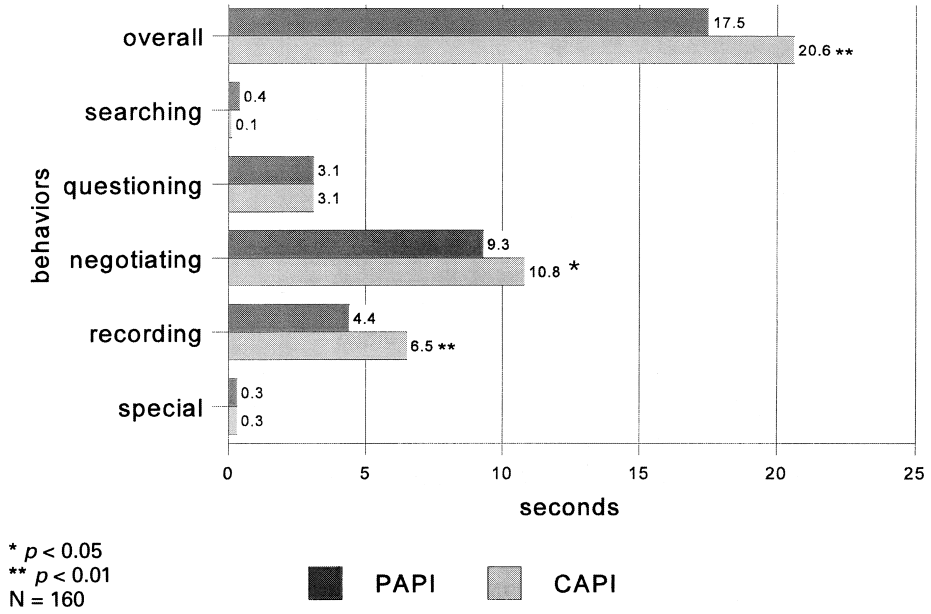

Fig. 6. Duration of "What are the names..." segment

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 nonverbal turns) take significantly longer on CAPI than on paper. This is mainly due to the degree of flexibility provided by each version. The computer-assisted instrument is designed as a multiple item screen combining several input fields on one screen; the input is expected in a certain order: first name, middle initial, and last name. If the respondent provides the last name first, the interviewer has two options. (1) He or she 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, he or she can record the last name and return to the first input field (by backing up twice) to type the first name and the middle initial. (2) The interviewer can ask the respondent to provide the information in the order expected by the computer system. (We suspect the second option is more likely to occur.) Whatever solution is chosen, it takes more time than the paper instrument, in which the interviewer simply writes the information in the appropriate blank spaces regardless of the order in which it is provided.

The second difference between the two questionnaire versions in terms of time is also related to the method of input. On 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. (Of course, the written entry in PAPI must later be keyed during data entry, but our focus here is on time during the interview.) This suggests that writing is faster than typing. One explanation for this difference may be that, when writing, interviewers make fewer or no errors, but it is the correction of errors when typing which takes longer. We find some support for this from the results of the usability test coding. In the paper version, two of the 44 names (4.5% of entries) required editing or correction, compared to 40 of the 131 names (or 30.5%) in CAPI.

Furthermore, the difference in terms of time for recording may be explained by the fact that interviewers are able to edit the paper questionnaire after the completion of the entire instrument based on a rough draft used during the course of the interview. Instead, the CAPI interview is no longer accessible to the interviewer once completed, so shortcuts in entering this information for later editing make little sense. Therefore, the interviewer must record all such information in the presence of the respondent. However, providing the interviewer with text editing functions and more flexible cursor movement on screens with multiple input fields might help speed up the recording of text 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 use of automated calculations and fills. In the socio-demographic portion of the NHIS instrument, the interviewer records the date of birth for all persons living in the household. In the paper questionnaire, he or she is required to use a paper form (similar to a calendar) to calculate the age from the given date of birth and the interview date. The same task is much 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?''

The results in Figure 7 show that the CAPI solution takes less than half of the time necessary to complete the same task on paper (4.6 vs. 10.5 seconds per person). This difference is mainly due to the fact that the interviewer spends almost four seconds per person working with the tool to compute the age and write down a two-digit number instead of pressing a single key for yes or no in the computer-assisted instrument.

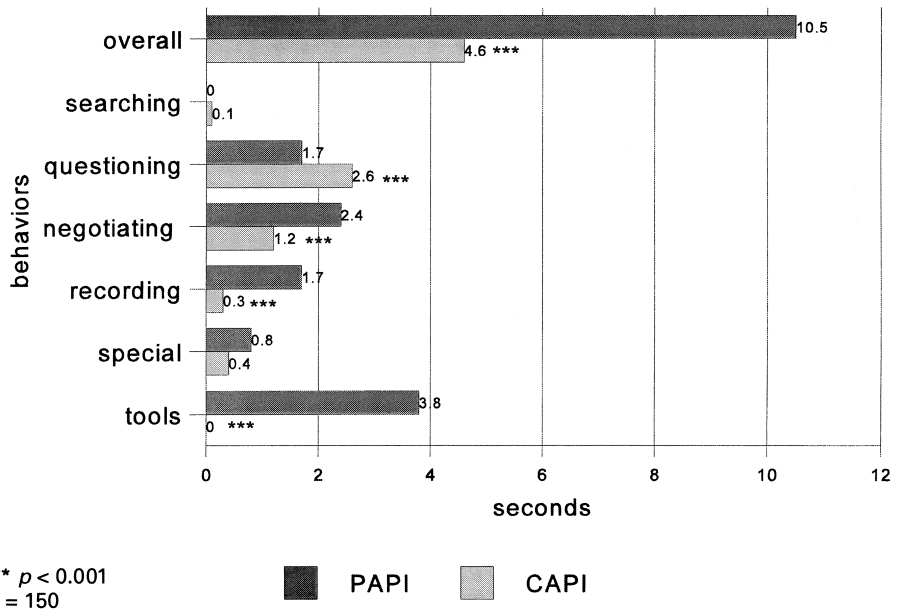

Fig. 7. Duration of ''Age'' item

Additionally, there is a greater amount of task-related negotiation in the paper version. As a result, the PAPI instrument requires more interviewer action and interaction between the interviewer and the respondent. This procedure is repeated for every household member, increasing the time difference between PAPI and CAPI for larger households.

We have shown earlier that recording text or character data using a keyboard takes more time than working with a pencil. Even though we cannot avoid this possible drawback of CAPI technology (as currently implemented), we can try to compensate for it in other ways. Making use of automated calculations and fills is one way to do this. Using questions like the age confirmation discussed here may have an additional positive influence on the interview. Making use of the computer's ability to store and compute information for reuse later in the interview, we can develop more intelligent instruments that reduce burden and support the interviewer's task.

3.5. *Direct technology effects*

So far we have discussed several differences between the PAPI and 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 compare a series of questions that has identical wording in both instruments and that applies the same question logic and expects the same responses or inputs from the interviewer. None of the design features previously discussed characterize this segment.

The segment we examine is 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 has not missed any persons. In order to verify the household

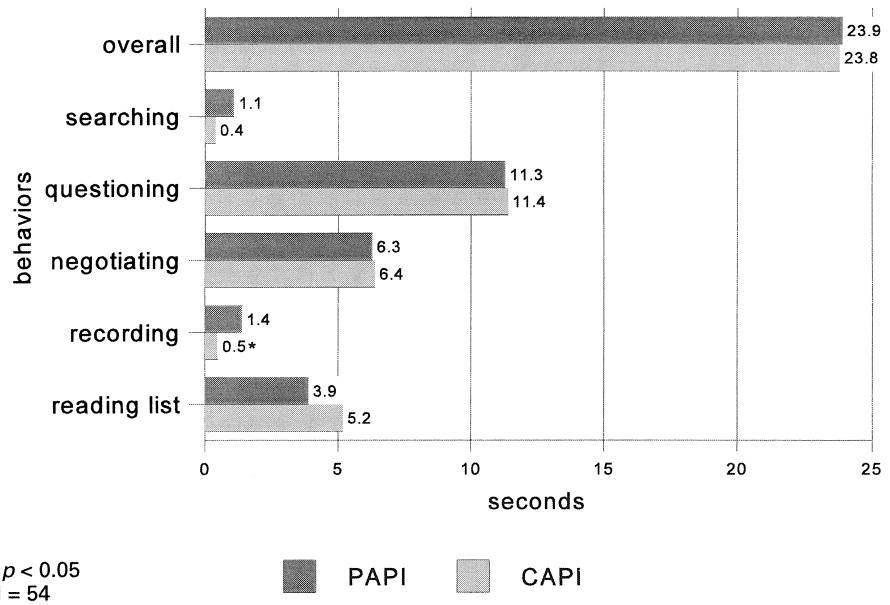

Fig. 8. Duration of "Missing persons" segment

composition the interviewer probes: "I have listed as living here ...," then reads the names of all persons living in the household, and then asks, "Have I missed any babies or small children? Any lodgers, boarders, or persons 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.

Examining the results in Figure 8, it appears to make no difference whether this household composition probe is administered using a computer or a paper instrument. The overall duration is identical, as is that for most of the individual components. The only difference that reaches 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 enter four single digits (1 = yes, 2 = no) and pressing [enter] in the computer-assisted instrument. Generally, though, this example suggests that the computer does not make much difference in terms of time if the compared sequence is not affected by other features introduced by the computer technology.

4. Discussion

The results reported here support the hypothesis that the introduction of the laptop computer into the interview situation does not necessarily affect the duration of the interview relative to paper. Still, the technology involved in a CAPI instrument comes into effect in several ways. (1) The overall interview duration can benefit from the system's ability to perform automatic calculations and make use of fills and other information previously recorded or drawn from external data files. (2) Typing text or character information using a keyboard takes more time than writing down comparable information with a pencil. (3) 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; however, this may depend on the number of entities in the loop. As a result of the combination of these factors, the CAPI instrument takes substantially more time to collect the same amount of information (an increase of 16.5%) for this part of the NHIS. In addition, other unobserved factors (such as interviewer performance) might contribute to the differences demonstrated for CAPI and PAPI instruments. We have also shown that when the sequence of questions is identical in design and execution in both computer and paper versions there are no time differences between the two technologies. These results suggest that whether the expected gains in time of CAI are achieved depends not so much on the technology employed, but more on how the instrument is designed and implemented.

It is apparent from these results that overall time measures of interview duration may mask many differences at the item level or at the level of specific interviewer and respondent behaviors. While time-consuming to do, the time measures we use here provide rich detail on the effects of instrument design choices. Item-level timers can be built into CAI instruments for pretesting of alternative designs, and can provide some indication of differences at the level of individual items (see Couper 1998). In addition, we used a coding scheme that contains data from a detailed assessment of individual turns and actions. This approach allows an in-depth understanding of interviewer-computer and interviewer-respondent interactions and their mutual influence.

It is important to mention that the effects discussed in this article are not attributable to differences in question wording. Most of the questions compared in this study apply the same question wording and the same response categories in the CAPI version as in the PAPI instrument. Furthermore, it is necessary to emphasise that we do not intend a software comparison. Some CAI products might make it easier to choose and implement appropriate screen designs. However, this article is on the effects of CAI screen design on the interview situation. It does not deal with the capabilities or features of certain software products.

So far, little is known about the effects of different design decisions on data quality. While there are no comparisons available for standard data quality indicators (e.g., item nonresponse), we do have some preliminary results from an analysis of data comparability. The findings suggest that different screen designs do not necessarily influence substantive data (Fuchs 1999).

In the next phase of this research we plan to address the effects of the screen design of the interviewer behavior when administering an item. As a first step we conducted a field experiment testing a set of identical questions applying four different design versions. Preliminary results indicate that the CAI screen design has a significant influence on the interviewer behavior and – in reaction to that – on the respondent behavior (Fuchs 1999). Further research on the influence of different design solutions on interviewer-respondent interaction is underway. Our goal is to explore the chain of effects from CAI screen design through interviewer behavior to respondent behavior, and finally to the result of this process in terms of data quality.

The results presented in this article as well as findings reported elsewhere (Hansen et al. 1997, 1998) contribute to a more general discussion about the usability of computer-assisted instruments. Our results suggest that technology influences the interview flow, the question logic, and the screen design as well as the interviewer's interaction with the instrument. Also, it appears that some CAI design solutions and screen layouts do not support interviewers in efficiently solving their tasks in the interview situation. This suggests that screen design is an important element of effective CAI development and use. Still, decisions on screen design in a CAPI instrument depend on many different factors, including the goals and specifications of the study, the software used, the programmer's preferences and skills, and the resources available for implementation. In addition, decisions on screen design often involve difficult trade-offs, considering interview coverage versus speed of interviewing and ease of navigation versus logical consistency. As a result, the expected effect of a proposed screen design in the interview situation is just one factor among others to be taken into consideration when planning design issues. However, this factor needs a careful assessment of usability to ensure that we benefit from technology advantages in terms of data quality without suffering from disadvantages caused by poorly designed screens and instrument features.

5. References

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

Baker, R.P., Bradburn, N., and Johnson, R. (1994). CAPI: An Experimental Evaluation. Proceedings of the Section on Survey Research Methods. American Statistical Association, Alexandria, 851–855.

Baker, R.P., Bradburn, N.M., and Johnson, R.A. (1995). Computer-Assisted Personal Interviewing: An Experimental Evaluation of Data Quality and Costs. Journal of Official Statistics, 11, 415–431.

Birkett, N.J. (1988). Computer-Aided Personal Interviewing: A New Technique for Data Collection in Epidemiologic Surveys. American Journal of Epidemiology, 127, 684–690.

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, eds. R.M. Groves, P.P. Biemer, L.E. Lyberg, J.T. Massey, W.L. Nicholls II, and J. Waksberg, New York: Wiley, 437–450.

Couper, M.P. (1998). Measuring Survey Quality in a CASIC Environment. Paper presented at the Joint Statistical Meetings of the American Statistical Association, Dallas, TX, August.

Fuchs, M. (1999). Screen Design and Interviewer Behavior. Paper Presented at the Annual Meeting of the American Association for Public Opinion Research, St. Petersburg, May.

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

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

Hansen, S.E., Couper, M.P., and Fuchs, M. (1998). Usability Evaluation of the NHIS Instrument. Paper presented at the Annual Meetings of the American Association for Public Opinion Research, St. Louis, MO, May.

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.

Lepkowski, J.M., Couper, M.P., Hansen, S.E., Landers, W., McGonagle, K.A., Schlegel, J., Wright, T., and Chevarley, F. (1998). CAPI Instrument Evaluation: Behavior Coding, Trace Files, and Usability Methods. Paper presented at the annual meetings of the American Association for Public Opinion Research, St. Louis, May.

Lynn, P. (1998). Data Collection Mode Effects on Responses to Attitudinal Questions. Journal of Official Statistics, 14, 1–14.

Lynn, P. and Purdon, S. (1994). Time-Series and Lap-Tops: The Change to Computer-Assisted Interviewing. In British Social Attitudes, eds. R. Jowell, J. Curtice, L. Brook, and D. Ahrendt. The 11th Report, Aldershot, 141–155.

Martin, J., O'Muircheartaigh, C., and Curtice, J. (1993). The Use of CAPI for Attitude Surveys: An Experimental Comparison with Traditional Methods. Journal of Official Statistics, 9, 641–662.

Müller, S. and Kesselmann, P. (1996). Akzeptanz von computergestützten Erhebungsverfahren. Ein empirischer Vergleich mit der traditionellen Fragebogentechnik. Marketing, Zeitschrift für Forschung und Praxis, 18, 191–202.

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

Waterton, J.J. and Duffy, J.C. (1984). A Comparison of Computer Interviewing Techniques and Traditional Methods for the Collection of Self-Report Alcohol Consumption Data in a Field Survey. International Statistical Review, 52, 173–182.

Received March 1999
Revised October 1999