DEVELOPMENT OF COMPUTER ASSISTED PERSONAL INTERVIEW FOR THE NATIONAL HEALTH INTERVIEW SURVEY

Stewart C. Rice, Jr., Robert A. Wright and Ben Rowe NCHS,
Stewart C. Rice, Jr., 3700 East-West Highway, Rm 2-44 Hyattsville, Maryland 20782

KEY WORDS: CAPI, NHIS, automated questionnaire, laptop, computer, CASS

INTRODUCTION: NCHS and the Bureau of the Census have been investigating the feasibility of automating NHIS data collection through Computer Assisted Personal Interviewing (CAPI). The use of CAPI and other computer assisted approaches to the collection of survey data offers significant potential for improving both the efficiency and quality of data collection and processing operations. In 1986, NCHS developed a research plan with the Bureau of the Census to create an NHIS CAPI system. The Bureau also planned to use this research to develop automation systems for some of the Bureau's data collection programs.

Automation of the NHIS household interview is expected to have a number of advantages. Among these advantages are an expected improvement in data quality through the edits of interview entries and control of the skip pattern of the interview. The CAPI application will not allow the interviewer to skip to an inappropriate section of the questionnaire since the skip pattern is part of the program. Automation of the interview also aids data preparation and data processing operations by reducing the clerical activities required to develop a computer data file. Therefore the efficiency of the NHIS program should be improved.

The basic purpose of the 1987 NHIS Automation Feasibility Study was to investigate the practical utility of using portable computers and a CAPI system to collect the NHIS data. No NHIS annual special topic supplements were included. The major objectives of the study were to assess interviewer and respondent acceptance and reaction to the use of a computer; to define hardware and software requirements for a complex, demographic health oriented survey; and to learn more about the type of interviewer training and support needed and how this differs from conventional training and support methods.

METHOD: The feasibility study was confined to two of the twelve Bureau of the Census regional offices (Charlotte and Chicago), with six interviewers in each region. In these regional offices, eight NHIS sample PSUs were selected. A total of 70 segments were selected to obtain approximately 450 interviews. The segments chosen had previously been designated for the NHIS but were not used because of a sample reduction in the NHIS survey design for 1986. They were selected to closely parallel a control sample of segments and housing units in the same PSUs that were currently in the NHIS. The segments included urban, suburban and rural areas. The field work was carried out between September 24, 1987 and December 12, 1987.

The study findings are based on a review of the interviewer and observer reports, interviewer training materials and debriefings.

TRAINING: All interviewers selected for the NHIS-CAPI Feasibility Test were previously trained and experienced with standard NHIS concepts and procedures. Therefore the NHIS-CAPI training was limited to how to conduct an NHIS-CAPI interview. the actual NHIS-CAPI application was used for all training demonstrations.

It has been estimated that well over 50% of the information presented to the trainees was needed to explain and circumvent software deficiencies and limitations in the CASS questionnaire. Even so, many software deficiencies were not covered, resulting in interviewer complaints of inadequate training. It is important to realize that most of the comments and suggestions for improving the training assume that it would be on the same NHIS-CAPI/CASS software. Improving, correcting or even replacing the software may be a better method for improving the training than developing a more intensive training package to more effectively cover software deficiencies and limitations.

In general, the Feasibility Test interviewers recommended intensifying the training to better prepare them for the field work. As one interviewer stated, "We don't like surprises and it's not good to feel uncomfortable in a sample household". Both situations were encountered during the feasibility test interviewing.

The interviewers also estimated that it would take less time to train new interviewers on an automated NHIS than a paper one because the computer did so much of the work.

PERFORMANCE IN THE FIELD: There were some inconvenient problems associated with the use of a laptop computer in the interview.

The interviewers had to spend more time paying attention to the screen than they do to the paper questionnaire. It was inconvenient to hand the flash cards to the respondent while holding the computer. They would sometimes set the computer down to hand the cards to the respondent. A result of these problems was that they were less able to give attention to sustaining rapport with the respondents.

Comments tended to be briefer in CAPI. Several interviewers said they entered comments in less detail because of a feeling they had to decrease the "dead" time between questions. A few said their typing skill governed the length of the comments.

The interviewers lost the "Big Picture" of the interview. They missed being able to see all of the names, addresses, conditions, etc. that were previously entered. If they thought a mistake had been made they could not verify it or they had to wait until later in the interview to see what happened. Sometimes they could not tell if a mistake was caused by an error in keying or a software error. There were many situations where they were not sure the entries were correct and could not do anything about it. They felt STRONGLY that they needed access to previously entered data during the actual interview.

The average interview took more time than a paper and pencil interview, but would probably take less time if the long pauses could be eliminated. The amount of time decreased as the interviewers became more familiar with the software. They did not like to have several weeks between assignments because they forgot some things during these free periods.

Respondents were mostly positive about the computer. There is no need for the interviewer to offer any explanation for the computer. It can be taken out and used without fanfare.

The CAPI application did eliminate many of the skip pattern errors commonly made by the interviewers. It eliminated the interviewer edit that is now required after the interviewer returns home.

An unexpected benefit was that the respondents could not see a questionnaire, so they did not realize how long the interview would last.

SOFTWARE: A survey instrument design work group was established in 1985 to evaluate available software packages for use in the design of automated questionnaires and the implementation of computer assisted interviewing. The work group evaluated several software systems, billed as suitable for CAPI, by attempting to "put up" a questionnaire which was a composite of the types of questions and procedures that appear in the Center's surveys. The software system selected was the Computer Assisted Survey System (CASS) by the University of Wisconsin.

The performance of the NHIS CAPI software was mixed when used in the field. The software rarely crashed or went into infinite loops. The problem reports included only nine cases in 387 interviews where the screen "froze up" or "locked". The group debriefing reported only occasional instances where a problem could be terminated only by turning the computer off. The software was relatively stable.

The interviewers liked the software's automatic branching between items and viewed this as one of CAPI's main benefits. They appreciated the system's automatic record keeping, especially its ability to store reported morbidity conditions by household member, sort them into a list, display them when required, and cycle through the appropriate condition questions for each. The interviewers also were generally pleased that CAPI freed them from editing the interview after they left the household.

Events that the NHIS interviewers reported as "software problems have a variety of interlinked sources: (1) the limitations of the CASS system; (2) the requirements of the NHIS; (3) errors in setting up the NHIS in CASS; (4) entry errors by the interviewers; (5) limitations in interviewer training; and (6) hardware constraints.

The CASS system has several limitations which make it less than ideal for a CAPI NHIS. The most important is its inability to permit interviewer backing across roster boundaries. When an error is made in a lead entry, the interviewer frequently is locked into an inappropriate branch (or skip path) from which there is no ready escape. Restrictions on backing also make it more difficult to record health events or conditions remembered after the basic questions covering them have been asked.

When an interviewer recognized that an error had been made, there were relatively few options, none optimal. The interviewer could accept the error and falsify the data until the interview returns to the next appropriate path. The interview could be aborted and started over from the beginning. Finally, the interviewer could abort the interview and leave the household with information which is both incorrect and incomplete.

All three options result in recording incorrect or falsified data, at least temporarily. There are several methods, all complex, that could be used to correct or note these types of errors. These complex methods appear to have been used only infrequently, however, and only by the most confident and technically proficient NHIS interviewers.

An additional limitation of the CASS system is the inconsistent forms of data entry it requires. For precoded items, the interviewer merely presses the precode number, say 2, and the system moves the next item. For numeric items, the interviewer enters the number and then presses the RETURN or ENTER key. For text entries, the interviewer enters the text and then must press the ENTER key twice to reach the next item. For multiple-answer list items, the interviewer uses the arrow keys to highlight the appropriate answer and presses ENTER once. However, to leave the item, the interviewer generally must enter a 0 and then press ENTER. These inconsistent entry methods may have contributed to the volume of interviewer entry errors.

The NHIS CAPI interview had frequent long delays. In weekly debriefing questionnaires, the interviewers were asked where the delays occurred. Most of them cited points in the interview where the computer stored data away for later use, compiled lists, completed calculations, or prepared summary displays. Specific points cited, which varied widely by interviewer, included calculation of ages from birth dates, the household roster display, hospital visits, restricted activities, doctors visits, and the race and ethnicity questions.

The most frequently mentioned delay, was at the end of the chronic condition list where the computer sorted conditions for a summary display by person. Test timings of this delay conducted in an office but with one of the interviewers' portable microcomputers, ranged from a minimum of 10 seconds to a high of 70 seconds. The length of the delay varied with the number of conditions reported for the person and the number of conditions in the household. Timings by number of conditions is shown in the following table.

Number of Conditions for Person

	1	2	3	4	5	6	7
Mean delay in seconds.	15.0	26.2	40.7	40.6	51.8	59.0	62.0
Cases timed.....	(5)	(6)	(3)	(5)	(4)	(2)	(4)

The affects of household size are suggested by the timings for persons with two conditions. The delay was 18.5 seconds in a one-person household, 24.0 seconds in a two-person household, 29.5 seconds in a three-person household, and 37.0 seconds in a four-person household.

The software problems not only proved inconvenient for the interviewing staff, they also appear to have resulted in the recording of incomplete or inaccurate health data in many cases. During the group debriefing, all twelve interviewers admitted that they had entered false or incorrect information at least occasionally to avoid or escape from software problems. One interviewer described the data as "very distressed at this point" and another estimated that only two of her 40 interviews were "good, that everything was all right."

The latter estimates may represent extreme views of the effects on data quality, for experiences apparently differed greatly among interviewers. A summary count was made of software problems which were likely to result in data errors. These problems were described in the interviewer problem reports and observer reports. Reports on the same case from both sources were accounted for. By this method a lower-bound estimate of the proportion of cases containing incomplete or incorrect data resulting from software problems may be set at about one-fifth of the

interviews. The true proportion may be considerably higher.

HARDWARE: The laptop computer selected for use in this study was the Toshiba T1100.

The Toshiba T1100 is IBM compatible. It has a 8086 processor at 4.77/7.16 MHZ. The weight of the machine is 10.26 lbs. Its Random Access Memory (RAM) size is 640K Bytes. There are two internal double sided, double density 3 1/2" diskette drives holding 720K Bytes of information each. The screen is a high resolution liquid crystal display (LCD) with 80 columns and 25 lines, 640 horizontal and 200 vertical pixels. The keyboard has 81 keys and is completely compatible with IBM software. It has a built-in rechargeable Nickel-Cadmium battery providing up to 8 hours of operating power. There is an AC adapter that allows operation off of standard house hold current. It also has an automobile power adapter. It has one serial port and an external diskette drive port. The dimensions of the computer are 12" x 11 3/4" x 2 1/2". The purchase price was $1998.60 (GSA).

The screen size was considered adequate by the interviewers. On the other hand, screen visibility/clarity was judged to be very poor in other than perfect lighting conditions. The majority of the interviews were conducted in other than perfect lighting conditions. Consequently, there were numerous complaints about eye strain, headaches, and loss of time adjusting and re-adjusting the screen angle, intensity, and contrast.

Keyboard layout and touch were considered adequate and acceptable by the interviewers.

The 10 lb. Toshiba 1100+ computer was considered by most of the interviewers to be too heavy. There were only a few complaints about the size of the computer. The interviewers found that respondents generally did not object to the use of their tables to support the computer. However, if a table was not available, the interviewers considered their laps as adequate support.

Interviews were sometimes conducted outside. Balancing the computer on some object other than a table or chair was reported by several of the interviewers as a problem. It was difficult to conduct interviews outdoors.

They reported a few problems with children or pets bothering the Toshiba. Everyone answered "No" when asked if anything was ever spilled on the machine. There were a number of complaints reported about the adequacy of the carrying case for the computer. Some questioned the ability of a soft case to protect the computer if dropped or banged. The major complaint reported was the inability to easily carry the required papers and ancillary items in a separate, sealable compartment outside of the case. Although not associated with a handling characteristic, several interviewers commented on the color of the carrying case and suggested that it be changed or have an easy to read Government identification. The interviewers indicated that the case gave them the mistaken appearance of being a representative of a religious group seeking contributions or converts. They felt that this situation resulted in some refusals. Internal storage (RAM) was adequate for the questionnaire program. External storage (floppy disk) was more than sufficient for the number of interviews conducted on any one day. Some interviewers, however, reported difficulties with the floppy disks. For example, some were confused as to which disk, program or data, should go in which disk drive. This situation resulted in either writing over the program disk or inability to "boot" the program. Also some of the data disks were not formatted which resulted in error messages and the inability to write data on the disks. It appears that the floppy disk problems were primarily due to interviewer error, lack of a good logistical control system, and/or lack of sufficient training or instructions for handling problems among regional office staff.

All but two of the interviewers relied almost exclusively on battery power to conduct their household interviews. The two interviewers who did not use battery power requested and received, in every instance, permission to plug the computer into a household wall outlet for power. Many of those interviewers who did use the battery exclusively or almost exclusively complained of short battery life. The typical battery life span reported was 4-5 hours. There were no reports of batteries not taking a charge once run down or of any interview being terminate because of battery failure. Approximately 50 percent of the interviewers used a car adaptor to charge the computer battery in the car. No problems with this method of charging were reported.

All of the interviewers and observers reported that the length of time needed for the automated questionnaire was too long. Greater speed could be obtained by either an improvement in the hardware speed, e.g., faster CPU processor, RAM disk, etc. or software design, e.g., reduce disk 1/0.

The interviewers were concerned about the ability to have the Toshiba repaired on a timely basis. It was suggested that each Regional Office have backup machines available in case the equipment breaks down.

Fourteen microcomputers were used during the study. Only one computer required maintenance. The "down" unit was out of service for one month. It spent three weeks in the Regional Office waiting to be sent out for service and one week to be serviced.

One interviewer indicated concern for the security of the computers because of their known value to the general public (newspaper and T.V. advertisement). The interviewer was concerned that the computers and/or interviewers could be a target for a thief. Given the number of hardware features and their potential conflicting requirements (e.g, increased power to provide greater screen readability requires more power which in turns translates into greater weight) the interviewers were asked to indicate the priority of the critical hardware features. They rated screen clarity as the highest priority, followed by the processing speed and the weight of the machine.

RECOMMENDATIONS: The recommendations for future CAPI efforts are based on interviewer reports, debriefings and observer reports.

The screen quality must be improved. The super twist LCD is not adequate. Electrolumenescent backlite LCD or gas plasma screens, for example, provide the needed readability for expected household interview lighting conditions.

The overall weight should be less than 10 lbs; the lighter the better.

A separate numeric key pad is preferred but not required.

A low battery indicator is required. When it becomes active, there should be enough power to allow down-loading for recovery.

The minimum RAM is 640K. A RAM disk capability, in addition to the 640K RAM, would be preferred.

Each interviewer should have a back-up battery with battery charger and car-charge adaptor. These

batteries should be rotated to insure maximum battery life and performance. Nicad batteries should be run down as low as possible before recharging.

Changing a battery should be very easy, e.g., slide cover.

The expected battery charge should last a minimum of 4 hours.

Processor speed must be increased to eliminate interviewer delay waiting for new screens (questions) to appear. This speed increase can be accomplished by increased CPU speeds (faster chip), removable ROM, RAM disk, and/or software improvements (e.g., minimal floppy disk I/O, compiled executable code).

Each Census Regional Office should have a trained support person capable of answering interviewer questions regarding the CAPI hardware and software.

Each Census Regional Office should have at least two spare computers for maintenance. When there is a computer breakdown, a spare should be used to replace the broken computer. A local maintenance Basic Purchasing Agreement (BPA) could be used to provide for computer repairs.

All data disks should be preformated for the interviewers or they should be instructed as to how to format a disk.

All program disks should be made non-write.

CONCLUSION: The Automated NHIS Feasibility Study provided a wealth of information concerning the use of a CAPI software system to collect the NHIS data. By noting and learning from the problems that occurred in the Feasibility Study, the ongoing CAPI development effort can continue to refine the methods and technology used in the NHIS automation effort.

The 1987 Automated National Health Interview Survey Feasibility Study has been followed with the implementation of CAPI in the NHIS 1988 data collection year. A 15 minute segment of the special supplement questions has been included in an automated form using the GRID-LITE portable laptop computer and a questionnaire programmed in DBXL programming language. CAPI research and development will continue throughout 1989 with current plans calling for the total automation of all NHIS special supplements in 1990.