SOME IMPACTS OF COMPUTER ASSISTED TELEPHONE INTERVIEWING ON SURVEY METHODS

Robert M. Groves, The University of Michigan
Marianne Berry, The University of Michigan
Nancy Mathiowetz, The University of Michigan

Developments in computer-assisted telephone interviewing are accelerating in the government, academic, and private sectors of survey research. The little existing documentation in this field concerns the nature of hardware/ software systems and descriptive accounts of survey administration using CATI. Few who work with CATI systems doubt that large potential gains in data quality and control over survey costs are possible with such systems. The advantages of CATI have been noted in several forums; we share the view that they are significant.

This paper, however, takes a more critical view of the technology and discusses some problems encountered by the researcher making the transition from paper questionnaire survey administration to computer-based interviewing. Such a transition requires certain adjustments in the activities of the survey researcher undertaking a data collection, adjustments which can be costly or productive of greater survey error. This paper focuses on two aspects of CATI which have proven, in our experience, to have some unforeseen consequences: first, the phenomenon which we have labelled segmentation, by which we mean limitations on one's ability to examine the survey instrument as a whole; and second, the standardization of survey activities which an on-line system compels. Both of these features, and their implications for various members of the data collection effort, are examined in turn below.

1. Segmentation of Tasks Within CATI

CATI in its most rudimentary form offers the researcher a storage and presentation medium for questionnaires. Since the user can see only one display at a time, this medium, although offering flexibility regarding the nature of the measurement process, restricts access to the entire questionnaire.

The restricted view of the questionnaire, a characteristic of on-line interviewing which we have labelled "segmentation," alters both the researcher's and the interviewer's tasks. Neither the questionnaire designer nor the interviewer can review the instrument in its entirety. Rather, interaction is limited to viewing sets of individual questions linked together according to the logic specified by the researcher. This transformation of the questionnaire from an object with a physical existence into a series of segments presented one-by-one over time, the "segmentation" of the questionnaire, tends to expand the role of the questionnaire designer, who under CATI has to make more explicit provisions to overcome the interviewer's restricted access to the questionnaire.

1.1 Effects on the Interviewer

Since interviewers can see only the single question displayed on the screen at any given time, it becomes one of the questionnaire designer's tasks to determine when and why an interviewer will need to consult another part of the instrument and to arrange how such scanning is to be done. With a paper questionnaire, no one has to explain to an interviewer how to flip backwards a page or two. But the designer of the on-line questionnaire must anticipate such needs and make explicit provisions for them.

When an interviewer wants to return to a previous question for clarification or to change an answer, he or she must be able to override the programmed logic of the questionnaire and indicate what question is desired. Just as CATI forces the researcher to standardize forward movement, similar decisions must be made concerning specifications for backward movement. When designing a machine analog for "turning back a page" several points should be considered.

1) Should an interviewer be able to review only the immediately preceding question or be able to review any prior response?

2) Once the interviewer has moved back in the questionnaire, how is a return to the point of departure easily facilitated?

3) If a change in one response affects branching from that question, should an interviewer be routed through the new branching?

Programming "special function" keys can alleviate both the problems of early breakoffs and backward movement. Early breakoffs create two problems for an interviewer utilizing an on-line system. First, an interviewer needs the ability to mark the interview as a break-off and skip out of the questionnaire to a section where comments can be recorded. Second, the interviewer who attempts to convert the partial interview needs to jump to the point at which the interview was terminated, perhaps reviewing the notes left by the previous interviewer. To aid backward movement one could program a key to return the interviewer to a point predetermined by the researcher, usually the last point prior to a branching sequence. Another key could allow the interviewer access to any previously seen screen. In the case of a breakoff, a third program would allow the interviewer out of the questionnaire to a special question where comments concerning the breakoff could be recorded.

The extent to which unusual movement is a problem will vary with the complexity of the questionnaire and capabilities of the system. Figure 1 illustrates some of these points. One can see from the flowchart that this questionnaire would be difficult to represent in a printed format. The flowchart represents alternative paths through part of the household listing and respondent selection portion of our first application. It should be noted that a flowchart offers the researcher one of the best means by which to review the questionnaire. The flowchart also clearly depicts how an interviewer could easily be unsure about the next screen that will be displayed. The number of unique paths is so large that familiarity with the line of questions is likely to be rather general.

Figure 1 also illustrates the difficulty of defining a backup route through the questionnaire. A large variety of paths could be taken to arrive at any one question. Interviewers who must go back to change an answer cannot be expected to remember how they got to a certain point. Either the program must "remember" and be able to retrace these steps, or else the interviewers must be able to specify the place to which they want to return.

Much of the complexity represented in Figure 1 is due to our system's lack of a "fill" capacity, which would have allowed us to insert the correct pronoun, for example, rather than creating multiple screens. Such a capacity would do some of the work of "remembering" previous answers for the interviewer. If, for example, the interviewer has to ask questions about each household member in turn, some systems will carry information from the household listing forward to the appropriate screens and will insert such information as the name, sex, and relationship of the household member into the text of the question itself. Such a capability renders the problem of contextual ignorance less acute, but it requires a predesignated planned program, and does not obviate the problem of the interviewer unexpectedly needing to consult some information from a previous screen.

1.2 Effects on Supervisors

The supervisors as well as the interviewers suffer the effects of segmentation. In off-line surveys supervisors use the height of stacks of coversheets of different statuses to estimate the amount of work that remains to be done. They can also note the speed with which interviewers process a group of coversheets in order to assess the optimal number of interviewers to use on the project. The assignment of numbers to interviewers can proceed with no action on the supervisor's part within CATI. CATI systems can be designed so that cases can move from one status to another without any reviewing required on the part of the supervisor. Toward the end of a survey period, when small numbers of cases remain to be called, a fixed definition of priority groups

will lead to the same numbers being called repeatedly. At this point in the survey new procedures should probably be implemented to guide the selection of the next numbers to be called.

Some approximation of the supervisor's review of all remaining coversheets can be easily made a feature of a CATI system. A procedure for reviewing sample number queues can be incorporated into the supervisor's tasks. This also facilitates a supervisor's knowledge of the number of cases in each queue. The call scheduling software must address the need of the supervisor to judge how many numbers of what type remain to be called. This requires some summary statistics available on-line for the supervisor. This problem is the analog of the inability of the interviewer to see sets of questions in the questionnaire.

1.3 Effects on Coding

In other stages of the survey process, segmentation creates a similar problem of contextual ignorance. Just as the interviewer's ability to see the entire questionnaire can be prevented by CATI systems, so too is the coder's ability to understand the flow of the interview conversation somewhat inhibited. By "coding" we mean the translation of answers to open-ended questions into numeric values. In terms of raw productivity it might be best to have coders move through the questionnaire from one open-ended question to the next, without seeing the intermediate questions.. This would have the coder focus entirely on the responses needing coding without obtaining information about other responses. We do not yet know how often such information changes a code from that expected from the answer alone, and some work is needed to measure that effect. Similarly, coders could be skipped around any comments that the interviewer might have placed on the questionnaire, and the marginal information those comments provide has not really be addressed in past work.

The presence of a data file on all cases in one location permits for the first time a practical alternative to the coding of all open-ended items for one questionnaire, followed by the same for the next questionnaire, etc. It is possible to code all questionnaires for one item; say, occupation for all units; then, another question for all units, until the entire questionnaire is completed. We don't know the error characteristics of such a procedure; one can guess that certain contextual information of the response may be lost, but the intensive analysis of one question may permit better codes for that question.

2. Standardization of Procedures Within CATI

Paper questionnaires often are accompanied by training manuals that specify the correct use of the questionnaire by the interviewers; in addition, skip instructions appear in the questionnaire so that interviewers can avoid asking irrelevant questions of subsets of

respondents. Despite these and other attempts to control the procedures used in the survey, off-line telephone surveys necessarily leave more activities to the judgment of supervisors and interviewers. CATI systems often force explicit and rigid specification of procedures that were formerly unspecified. This feature of CATI has been promoted as one its chief benefits. Researchers must learn, however, that this benefit has its costs, both in terms of presurvey activities and errors from incorrect specification. Some of the reasons for standardization of procedures are linked to the CATI technology itself. For example, measurements used in CATI are limited to those that can be presented on a terminal screen that usually contains 24 lines of 80 characters each; having a programmed solution to assigning sample numbers to interviewers and monitoring their status forces less flexibility than was available when paper coversheets were used to identify sample numbers. Other causes for standardization occur not because of the technology alone, but rather as attempts to ease the effects of segmentation. For example, some systems use special screens to present data previously obtained from the respondent to aid the interviewer in her task (e.g., household composition). The investigator must assess what information is most often needed and present that subset despite the fact that different situations might call for different subsets. This section discusses some of these features of CATI systems and comments on their effects on the researcher and the interviewer.

2.1 The Format of CATI Screens

The most obvious example of standardization is the display of questions within the constraints of a terminal screen. Figure 2 presents a terminal screen as formatted in the PILOT, the first SRC CATI system. The SRC system utilized two separate answer spaces, one designated for numeric, precoded responses and another for open-end or verbatim responses. While some systems do not make this distinction by the use of a predesignated physical space, as ours did, some delimiter is usually used to distinguish precoded answers from those that need to be coded. When space is the delimiter, the cursor must be moved to the appropriate area of the screen; in PILOT the researcher had the ability to program the initial position of the cursor. For many questions, which required only a numeric response or a response to an open-ended question, there was little problem in deciding the initial position of the cursor. On several screens, however, we required both a numeric and a verbatim response. We had then to decide the order in which the interviewer was likely to use the two spaces (would the interviewer need to go directly to the open space to record spontaneous comments, or would the numeric answer be recorded first?) and initialize the cursor accordingly. Moreover, having set up a pattern to handle the typical respondent, we had to consider whether the interviewers had a sufficient range of behaviors available to them with which to handle an atypical interaction.

For numeric responses, we mentioned that the interviewer chooses an answer from a list of codes displayed on the screen. To end the survey period with no wild codes in the data, to have an automatic checking system that stops the interviewer from entering a wild or inconsistent value, seems like a great advantage of CATI, but it has its costs. As the questionnaire becomes more complex, the number of paths through the questionnaire an interviewer can use increases substantially. The questionnaire designer must consider every possible path through the interview in order to garner exhaustive lists of valid and invalid codes. Moreover, some provision must be made for the interviewer's behavior subsequent to receiving an invalid code message --- whether to go back to resolve the inconsistency or to be required to change the current code before proceeding.

Many of the rather rigid procedures of a CATI are the result of attempts to deal with the lack of information available to the interviewer at any particular moment during an interview. For example, lines at the top of the screen could be programmed to contain relevant information for either the interviewer's or monitor/supervisor's utilization. Information concerning the household composition, the length of the interview, or identification of the current section of the questionnaire can be presented to aid the interviewers in their work. A monitor or supervisor viewing copies of interviewers' screens on a monitoring terminal will be assisted by identification of which interviewer is using the terminal, what telephone line is being used, and what telephone number has been reached.

2.2 The Effect of CATI Procedures on Interviewer Selection

The inflexible demands of the technology placed on the interviewers forces new concerns for the selection and training of interviewers. Decisions concerning the screening of interviewers, whether to train solely using the on-line system or whether to use a "hard copy" in training, and determining the extent to which an interviewer needs to understand the system logic are among the questions to be resolved.

The first aspect of interviewer training affected by use of a CATI system comes during applicant screening. Legibility of handwriting is no longer a relevant criterion for assessing interviewers; the interviewers now need to be evaluated on their ability to type verbatim responses rapidly and accurately. Time for training interviewers to follow skip patterns is greatly reduced, but is replaced with the need to include time for reducing machine-related anxiety. The initial instruction, which includes information about signing on the system, forward and backward movement, and learning to use programmed keys, etc., is now often taught in approximately six to ten hours. Practice time, in which trainees interview one

another and then begin calling non-sample numbers, varies significantly among individuals, due in part to the time needed to reduce machine intimidation for some. Many of the college students we trained were familiar with either CRT terminals or had interacted with computers often enough to feel somewhat relaxed in using the on-line system. For others, often individuals who had been out of the job market for an extended length of time, one or two extra sessions during which special attention was given were adequate to reassure them of both their ability to conduct a machine assisted interview successfully and of their inability to "break" the system. We found that three four-hour sessions allowed adequate practice time for everyone, with the majority of interviewers needing only two sessions to feel comfortable with the system.

2.3 Automation of Sample Administration

Software to permit the automatic assignment of sample numbers to be called is often part of a CATI system. This software offers a good example of the unanticipated effects of standardization. In this case, the CATI rules for assignment of coversheets to interviewers and scheduling of callbacks may have been implemented before a complete understanding existed of the procedures of sample administration without computer assistance. The administration of sample numbers is one area where rigid procedures are generally not implemented, and instead callbacks are made on numbers in some fashion that may vary across interviewers and supervisors. The initial attempt at constructing a procedure for call scheduling was one that tried to quantify the large collection of decision rules that supervisors now use to determine which coversheets to give to interviewers needing new numbers to dial. This process includes sorting cases into time zones so that numbers that are not called too early or too late, creating categories for numbers that have received many unsuccessful dialings already, separating appointment cases from others and keeping records on when they should be called, checking on the number of coversheets, etc.

There are several observations that can be made rather quickly when faced with the vast array of decisions that the supervisors make about distribution of sample numbers. At this time, and especially at the first implementation of the procedure using machine guided selection of dialings, we cannot fully quantify the large number of decisions that are made. It is difficult to determine an orderly machine algorithm for all numbers. One category of sample numbers that causes problems is the group which require appointments with respondents before completion. Some appointments can be quite specific, requesting a call on a certain date and time, while others may be general (e.g., any Wednesday except next week). Some handling of these cases by interviewers or supervisors seems inevitable. One scheme attempted at SRC yielded poor results and offers

a general lesson. We asked interviewers to document on paper slips appointments that they could not keep so that supervisors could give them to interviewers working the specific shift. Interviewers made errors on the slips, the slips were lost, supervisors sorted them improperly. The lesson we think is that one should never try to move in and out of machine readable form. The system should have provided for documentation of the appointment cases in a fashion that allowed supervisors to have control over them without having them lost. Hard copy documentation is desirable only when there is no more need for machine actions on the case.

3. Summary and Conclusions

Both segmentation and standardization tend to have the same net effect, which is the expansion of the activities of questionnaire design and of the responsibilities of the researcher. This concentration of decision-making into the early stages of the questionnaire is due to several factors:

1). Segmentation creates contextual ignorance. The researcher must therefore anticipate situations in which an interviewer or coder will need access to information other than what is on the currently displayed screen, and make provisions for such access.

2). The "invisibility" of the questionnaire itself requires the generation of new types of information to replace what was formerly visually ascertainable. For example, summary statistics must be provided for supervisors who previously could judge the progress of the study by monitoring the accumulation of coversheets.

3). The necessity of standardizing and specifying all activities into a programmed logic unearths a plethora of decisions which were hitherto made on an ad hoc or idiosyncratic basis. A particularly illuminating instance was our difficulty in quantifying an algorithm for assigning appointments to interviewers.

This distillation of decision-making into the early stages of design is both an advantage and a disadvantage of the transition to CATI. The researcher clearly has greater control over the data collection process. Such control can prove a two-edged sword, however, if aspects of the survey that were successfully accomplished previously, albeit not according to rigid procedures, are subjected to premature or unnecessarily inflexible standardization. By the same token, there is great flexibility offered researchers in the number and kind of quality checks entered into the CATI questionnaire at the time of construction. Those that do not take advantage of these powers, or who do so without sufficient forethought, will be able to do poor work within the system. The promise of a higher quality of work offered by CATI is an opportunity, not a guarantee.

Case =   Family # =   Person # =   Cond. # =   Interviewer=   Wats=

RF7.   In this household I'd like to interview a male adult.
        Since you mentioned that there is only one man in
        the household, may I (speak with him / interview you)?

        1.   YES

        5.   NO

****

****
***   ***
***   ***
***   ***
***   ***
***   ***
***   ***
****

Figure 2.   Example of CATI Display

Figure 1. Flowchart for Segment of a Questionnaire

STRT

1,2
3,4

SEX

STRT=2   STRT=1

SEX

STRT=3   STRT=4

R1   R1P

1-4   1-4

5   5

RB1   5   TERM

R2   R2P

RNIP   1   1   RNI

5   5

R3   R3P

1   5

R3A   5   TERM   RN3Q   5   TERM   5   RN3A

TERM   5   R3A

1   5

1   5

R3AP   5   TERM

TERM   5   R3B   R3BP   5   TERM   RN3R   5   TERM   5   RN3B

1   1   1   1

R3C   R3CP   RN3S   RN3C

RB2   5   TERM

1
STRT=2   STRT=1

R1R   R1RP

1-4   1-4

R2R

R3R

SEX=1
SEX1=1   SEX=2
SEX1=2

BRNH

TERM

RM5   RF5

1   0   0   1   2-6

RM6   2-6   RM7   2-6   RM8   0   HELP   0   RF6   RF9   RF11

0   1   0   1   2-6

1   1   2-6

RM1   RM10   RM11   RM12   RF7   RF8   RF10   RF12

5   1   5   1   1   5   5

RMIA   RMZA   RF7A   RF8A

TERM   TERM   TERM   TERM

REQR

524