Operational and Quality Issues with Computer Assisted Interviewing for ADAM

David Cantor
Westat and
Joint Program in Survey Methodology, University of Maryland

Paper for OJP Offender Drug Abuse Monitoring Program, BJS-NIJ Expert Topic Meeting, May 27, 2010

5/24/2010

In this paper we discuss the use of computer assisted interviewing technologies for the Arrestee Drug Abuse Monitoring (ADAM) survey. Currently, ADAM is collected using a paper and pencil instrument (PAPI). The discussion below provides an overview of the advantages and disadvantages of using a computer assisted methodology with respect to five areas: 1) cost, 2) data quality, 3) timeliness, 4) survey management and 5) logistics.

The computer assisted interviewing (CAI) technologies that are seem most relevant to for ADAM are computer assisted personal interviewing (CAPI) and computerized self administered interview. CAPI applications involve an interviewer administering the survey from a portable computer of some type (laptop, tablet computer; PDA). The self-administered interviews could take several forms. A computer assisted self interview (CASI) has the respondent reading the question from the computer screen and entering answers directly into the computer. An audio-CASI (ACASI) provides headphones to the respondent who is able to listen to recordings which read the questions as well. Interactive voice response (IVR) administers the interview over the telephone using a computer program that administers the questions using recorded voices.

Costs

For an ongoing program like ADAM, there are cost savings and additional costs associated with CAI. One source of cost savings is the reduction of data processing costs. The use of a computer eliminates the need for data entry (either manual or scanning). The use of CAI will significantly reduce, but not totally eliminate, data editing after the data are captured. Editing is reduced because of the automation of the skip patterns. The ADAM instrument has a significant number of skip patterns, including the use of the event history calendar (EHC), which inevitably leads to data that need review and alignment. Related to this is the elimination of the need for interviewers to conduct any post-interview edits. In most PAPI surveys, the interviewers need to spend time reviewing their markings on the questionnaire after the interview. This activity is minimized with CAI. A second source of savings is the elimination of the need to mail hardcopy questionnaires. Interviewers will electronically transmit data to the home office on a regular (probably daily) basis.

There are two major sources of additional expenditures associated with CAI. One is the investment in the hardware. This includes the initial purchase and continued maintenance of the machines. Each interviewer and supervisor needs their own machine. There has to be a significant number of backup machines that can replace machines that have problems during the field period. Along with the machines, there is a need for systems personnel to support the field staff. A second cost is the specification, programming and testing of the program. This is a one-time cost incurred whenever a new set of questions are in the field. A third additional expenditure is training staff to use the machines (e.g., use of sample management system; navigation around the instrument; transmission of data after interview). This is an additional set

of modules to the training program that would not be necessary if a paper instrument were being used.

There are no precise figures on the relative costs of CAI vs PAPI surveys. The comparisons are difficult to make because conversion to CAI typically introduces new capabilities and complexities that can be handled by the technology. An important consideration for an ongoing program, like ADAM, is the amortization of the fixed development and hardware costs. CAPI offers much clearer cost savings if machines and computer programs are used for extended periods of time. A second important offset to these additional expenditures is the amount of data that is being collected. As the number of interviews increases, the greater the savings on data capture and editing costs. When CAPI was first being adopted by different survey programs, several organizations reported the reduction in costs as being an important reason for making the change (Martin and Manners, 1995; Rothschild and Wilson, 1987; Baker, Bradburn and Johnston, 1995). Whether these savings would apply for ADAM depends on the overall sample size and the extent of savings that would be realized by data capture, reduction in transmission and editing.

Data Quality

There are a number of features related to CAI that have effects on data quality.

Control over the interviewing process

There are at least three features of CAI that should improve data quality by increasing control over the interviewing process. Relative to an interviewer administered paper survey (as ADAM is currently), CAI increases control over the sequence the questions are asked. With a paper survey, the interviewer has the ability to move through the instrument in any order. This is generally not viewed as a positive influence on quality because interviewers can then take shortcuts by not asking all questions. This can be particularly important if the questions are a series of items asking about specific behaviors, such as in several item-sets on the ADAM II instrument (e.g., S1, S4, S10, S13, S16, S19, MU36 a - n)). One purpose of using lists like this is to prevent respondents' prematurely ruling out the occurrence of an event ("failure of metamemory"). However, in the interest of completing the interview, interviewers might be inclined to skip items if the respondent reports not engaging in any of the behaviors before they are actually asked the questions. Evidence that automation may have effects like this was found in tests related to the National Crime Victimization Survey (Hubble and Wilder, 1988).

Imposing structure on a CAI may negatively affect data quality when it is important to give the interviewer flexibility to navigate questions. The event history calendar is a procedure that does require this type of flexibility. The interviewer is trained to probe based on what the

respondent might say to different items. It may also be the case that when working with arrestees situations may arise when it is difficult to work in a specific question order. Interviewers may need the flexibility to skip around the questionnaire.

This raises the question of whether it is possible to use an event history calendar (EHC) with a CAI application. As noted above, the EHC is a relatively unstructured protocol. In addition, it relies on filling out a hardcopy version of a calendar. There are now a number of applications that have used a computerized version of the EHC. For example, Belli, et al. (2007) found an EHC administered as part of a computer assisted telephone interview was very effective in collecting information when compared to a more standard set of question lists. The Census Bureau is now in the process of implementing a CAPI version of the EHC for the Survey of Income and Program Participation. What remains relatively unknown is how computerization affects the effectiveness of the EHC vis-à-vis a paper version of the same protocol.

The structure imposed by the CAI has the effect of eliminating the need for interviewers to manually navigate skip patterns on the questionnaire. The automation of skip patterns reduces the amount of missing data that results from interviewers failing to follow some of the skips. Theoretically this frees up the interviewer to concentrate on working with the respondent. This advantage tends to be more important at the beginning of the field period, when interviewers are learning the skips. For the current ADAM II instrument, this learning curve could be significant, given the dependencies many of the questions have on prior answers.

A third feature of CAI that affects control over the interviewing process is the availability of paradata that can be collected as part of the case management and survey interview. Paradata refers to information that is collected about the data collection process. This might include, for example, the number of times the interviewer attempted to complete an interview with a particular respondent, the amount of time the survey (or particular sections) took to administer and even the keystrokes interviewers used when entering the information. Timing information can be especially helpful because it provides a measure of how much time interviewers are spending on particular items/questions. If they are rushing through certain sections, timings can provide a window into this. Similarly, keystroke files can provide some indication of how often interviewers have to back up, erase or re-do answers. This can help monitor the performance of particular items on the questionnaire.

Collecting sensitive information

The use of CAI introduces the possibility of using a self-administered questionnaire, such as CASI, ACASI or perhaps even IVR. Self-administration has been found to elicit better quality data for sensitive or illegal behavior (Tourangeau and Smith, 1998). Using a self-

administered paper survey may not be possible, given the skip patterns involved on the questionnaire. However, it is not clear the effects of self-administration generalize to the unique situation of interviewing booked arrestees. The studies that have found self-administration is optimal have been conducted with general population samples. Offender samples, in general, may not have the same inhibitions related to reporting drug use or other criminal behavior. For example, the original offender studies conducted by RAND found, if anything, that some offenders tend to overreport their criminal activities (Blumstein, et al., 1986). There is the immediate legal threat related to their arrest which may inhibit reporting. If a self-administered survey were to be used for ADAM, some type of experimental test would be needed to assess it's effects on data quality.

Many of the analytic uses of the ADAM interview are related to the details associated collected about drug use and offending. For example, the ADAM interview collects information on the types of drugs offenders had been using prior to arrest, how they were using them, dependence on drugs/alcohol and how the drugs were obtained. Respondents may be more willing to report these details with a self-administered questionnaire, although it is not clear from existing research. These items are amenable to a CASI or ACASI application. The program could guide the respondent through relevant skip patterns, which would be difficult in a self-administered paper questionnaire. However, it is questionable if the EHC could be done as a self-administered application. It may be possible, again using the computer's routing and visual features. However, this would require significant development. At least one recent attempt to conduct a EHC with a paper self-administered version would indicate that it would still require some intervention by an interviewer to assist in the process (Cotugno, 2010).

Online edits

The accuracy of the data for CAI applications has been found to be comparable to manual key-entry or scanning (Dielman and Couper, 1995; Lepkowski, et al., 1998). A CAI application offers a way to check the plausibility of values during the interview. Plausibility includes whether data are within realistic ranges and whether there is consistency between questions. When responses do not seem acceptable, the program can alert the respondent and ask for either clarification or a corrected value. This can be done in either an interviewer or a self-administered mode.

Timeliness

In terms of timeliness, a CAI application requires more up front planning and testing. The program has to be specified by designers, the programmers have to implement those specifications and the programs have to be tested. This lead time increases as the complexity of the program increases, as well as when an audio component is involved (i.e., ACASI, IVR). The opposite is the case for data production and file creation. The use of CAI greatly increases the speed with which data-sets can be created and analyzed. It is even possible to analyze data, at least in its raw form, within days of receiving it from the field. This provides a capability to track interviewer performance on the questionnaire (e.g., using the para-data referred to above), as well as tracking results related to the questionnaire items.

This capability requires careful planning. Survey designers have to decide early on the specifications for the program, which would be translated for programming. In addition, there needs to be careful planning of of the transmission protocols used by the interviewers. The transmission needs to be relatively easy to implement, support has to be provided to deal with problems and, most importantly, careful attention to the security of the transfer has to be considered.

Survey Management

CAI applications have management systems that can administer sample cases to interviewers. This is a powerful tool for tracking the disposition of particular cases. For example, electronic records can be maintained to keep track of contact attempts (including time and date) and detailed disposition codes. This also makes it relatively easy to transfer cases to different interviewers. For ADAM, this capability would have to be integrated into the sampling methodology. Since the sample is not defined until just prior to interviewing, it would be necessary to have a procedure that enters either the frame or the actual sample case on a flow basis. This type of updating has been done when sampling and interviewing youth in residential placement (Sedlak, 2008; Beck, et al., 2010). However, in these cases, the ability to collect the sample data can be planned a day in advance. This may not be possible, if the ADAM sample needs to be drawn just prior to interviewing.

ADAM also requires data be collected from administrative files which is used to fill out the face sheet. The collection of these data could also be computerized. Interviewers would enter the information into the computer, rather than filling out the face sheet by hand. This could then be incorporated as part of the interview record or kept separate if that is necessary for confidentiality reasons.

A final capability of a CAI application is the ability to distribute survey instruments electronically to interviewers. This might be in the form of an update to the programmed instrument or the addition of a new module.

Survey Logistics

The application of CAI to ADAM is subject to the relatively unpredictable, and sometimes chaotic, atmosphere of a booking facility. The above discussion has already mentioned the possible complication with respect to drawing the sample. Facilities will vary by how they will provide the sample and it is not clear how this might fit into using a computer application. A second question is whether there is the physical space needed to use a laptop. Is there somewhere that the computer can be set up? Lightweight laptops, as well as tablet computers, can be designed to be used without requiring a flat surface (e.g., on the doorstep of a house). This would accommodate doing the survey, in the worst case, while standing in front of the holding cell. A related question is access to a power source. It is preferable to be able to draw power from an electric outlet, rather than use the computer battery. This eliminates the possibility that the computer will run out of power. The worst case scenario would be to use the battery, but in most cases it should be possible to use extension cords, strategically placed and hidden, to accommodate most facilities.

Security of the machines is also a concern that facility administrators, and project staff, voice when using computers around offenders. The extreme concern is that the respondent will intentionally break the machine. Less extreme concerns relate to the security of the machines when they are not being used by staff.

To our knowledge, there is very little extant experience with using CAI in the context of interviewing arrestees in a booking facility. The one application we are aware of is the Substance Abuse Need for Treatment among Arrestees (SANTA), a study sponsored by CSAT. This study used CAPI to interview arrestees in booking facilities. We do not know many of the details of this implementation, but it would be useful to follow-up with the sponsors to get more details, if CAI applications are being considered for ADAM. It is the case that CAI applications have been used when conducting group administrations among juveniles in residential placement (Sedlak, 2008) as well as individual interviews with juveniles in residential placement and adult prisoners (Beck, et al., 2008; 2010). These applications faced some of the same challenges as noted above, such as possible damage to equipment, security, sampling updates and power sources. All of these were successfully overcome through customization of the computer systems, as well as creative solutions to accommodate the physical layout of the buildings.

References

Baker, R.P., Bradburn, N.M. and R. Johnson (1995) "Computer Assisted personal Interviewing: An Experimental Evaluation of Data Quality and Survey Costs." *Journal of Official Statistics* 11: 415-434.

Beck, A. and P. Harrison (2008) "Sexual Victimization in Local Jails Reported by Inmates, 2007." Bureau of Justice Statistics, U.S. Department of Justice, NCJ 221946

Beck, A., Harrison, P. and P. Guerino (2010) "Sexual Victimization in Juvenile Facilities Reported by Juveniles, 2008-09." Bureau of Justice Statistics, U.S. Department of Justice, NCJ 228416.

Belli, R. F., Smith, L.M., Andreski, P.M. and S. Agrawal (2007). "Methodological Comparisons between CATI Event History Calendar and Standardized Conventional Questionnaire Instruments." *Public Opinion Quarterly* 71:603–22.

Blumstein, A., Cohen, J., Roth, J.A. and C.A. Visher (eds) *Criminal Careers and "Career Criminals.* Washington D.C.: National Academy Press.

Cotugno, M.Z. (2010) "Creating a Self-Administered Questionnaire." Poster presented at the 2010 annual convention of the American Association for Public Opinion Research, Chicago, IL, May 12-15.

Dielman, L. and M. Couper (1995) "Data Quality in CAPI Surveys: Keyig Errors." *Journal of Official Statistics*, 11:141-146.

Hubble, D. and B.E. Wilder (1988). Preliminary results from the National Crime Survey CATI Experiment. *Proceedings of the American Statistical Association: Survey Methods Section*, New Orleans, LA.

Lepkowski, J., Sadosky, S. and P. Weiss (1998) "Mode, Behavior, and Data Recording Error" pp. 367-388 in Couper, M., Baker, R., Bethlehem, J., Clark, C., Martin, J., Nicholls II, W. and J. O'Reilly (eds) *Compuer Assisted Survey Information Collection.* New York: Wiley.

Martin, J.and T. Manners (1995) "Computer Assisted Personal Interviewing in Survey Research." Pp. 51-72 in R.M. Lee (ed) *Information Technology for the Social Scientist.* London: UCL Press.

Rothschild, B.B. and L.B. Wilson (1988) "Nationwide Food Consumption Survey 1987: A Landmark Person Interview Surveys Using Laptop Computers." *Proceedings of the Fourth Annual Research Conference*, U.S. Bureau of the Census, pp. 341-356.

Sedlak, A. (2008) "Introduction to the Survey of Youth in Residential Placement." Juvenile Justice Bulletin, Office of Juvenile Justice and Delinquency Prevention, U.S. Department of Justice.

Tourangeau, R., and T.W. Smith. 1998. Collecting sensitive information with different modes of data collection. In *Computer Assisted Survey Information Collection,* edited by M. Couper, R.P. Baker, J. Bethlehem, C. Clark, J. Martin, W.L. Nicholls II, and J.M. O'Reilly. New York: John Wiley & Sons, Inc., pp. 431–454.