

DISCUSSION

William L. Nicholls II, U.S. Census Bureau
Field Division, Census Bureau, Dept. of Commerce, Washington D.C. 20233

We have heard four contributed papers on computer assisted data collection -- or at least four papers that mention CATI or CAPI in their titles or first paragraphs. They range from a speculative overview of issues in CAPI development to summaries of complex research programs designed to assess the effects of CATI in specific surveys. I will begin their discussion with the two papers on CATI by my Census Bureau colleagues.¹

Shoemaker, Bushery, and Cahoon have summarized a team research effort by Census Bureau staff undertaken to assess the effects of centralized telephone interviewing and CATI on the Current Population Survey. As a continuing, although peripheral, member of that team, I find it difficult to find new points of criticism or praise as that research draws to a close. Instead, I will offer a few comments placing that research in a broader perspective.

First, this was a major research program, extending over several years, costing more than a million dollars, and drawing on the research staff of several Census Bureau divisions (U.S. Census Bureau, 1988B). Measured by sample size, staff effort, or cost, this was the largest evaluation of CATI thus far undertaken.

Second, because this research focussed very specifically on one survey, the CPS, and contrasted a mixed-mode test methodology with mixed-mode traditional methods, it provides few conclusions about CATI generalizable to other surveys. Its primary function was to ensure that the introduction of CATI did not produce any unanticipated, uncontrolled, or detrimental method effects on CPS estimates of the population and labor force.

Third, it concluded that the test methods did not produce any major detrimental effects, or at least none that could not be managed, controlled, or (if necessary) terminated by halting the closely monitored implementation process. As a result, partial introduction of the test methodology using CATI in the CPS was allowed to proceed.

Fourth, the CATI interviews in this test were designed to parallel paper-and-pencil methods as closely as possible. Redesign of the CPS questionnaire to take advantage of the special capabilities of computer-assisted interviewing is scheduled next with the introduction of CAPI for personal interviews (U.S. Census Bureau, 1988A). If appropriate funding is forthcoming, an even larger research effort will follow to assess their combined effects.

This is not the end of CATI evaluative at the Census Bureau. At last year's meetings, Hubble and Wilder (1988) analyzed the effects of CATI on estimates of crime victimization in the National Crime Survey. A study of the effects of CATI on the American Housing Survey are in progress and may produce some intriguing results by next year (Schwanz *et al.*, 1988). And in the final paper of this session, Durant and Vitrano described another extensive research program to assess the use of centralized telephone interviewing and

CATI in the Point of Purchase Survey. The PPS CATI evaluation is just now yielding first results. Thus, CATI evaluation remains a major activity at the Census Bureau. It is not, I believe, that the Census Bureau mistrusts CATI *per se*. Rather, the Census Bureau traditionally evaluates new methods very carefully when their introduction could disrupt important data series.

Much of the Durant and Vitrano paper should perhaps be regarded as a "preview of coming attractions." They are not able to tell us much this year beyond two very familiar findings: (1) that commercial samples of listed telephone numbers yield higher response rates than RDD samples; and (2) that the response rates of clustered RDD are higher than those of simple RDD. Since both results were first published by others more than a decade ago and have been repeatedly replicated since, these conclusions are not surprising (Waksberg, 1978; Groves and Kahn, 1979; Lepkowski, 1988).

The Durant and Vitrano paper would have been far more interesting had it included some details on the coverage and undercoverage of the list frames by geographic area and provided estimates of survey costs under varying dual-frame designs. The first of these topics, I understand, is being addressed by Clyde Tucker in a separate session of these meetings. The second topic remains an objective for the future. This year the first results of this potentially valuable research program were in my opinion stretched across too many papers.

The Shulman and Adams paper, "Mail Questionnaires: An Alternative to Telephone Surveys," does not deal with CATI or CAPI except as rejected alternatives. How fast things change. Only five years ago, in November 1984, the OMB Statistical Policy Office issued Working Paper 12 recommending that Federal agencies reduce their overwhelming reliance on mail questionnaires and consider use of modern telephone methods. The authors of that report would undoubtedly be amazed at their apparent effectiveness. We now have a paper suggesting that things have already gone too far; that we should reverse directions and consider mail questionnaires as an alternative to telephone interviews. I'm being frivolous, of course, just as Shulman and Adams were somewhat frivolous in choosing a title for their otherwise commendable paper. The mail questionnaire is still the overwhelming choice for Federal data collection.

Shulman and Adams present a very persuasive case for the use of mail questionnaires rather than telephone interviews as the primary data collection method for the Pregnancy Risk Assessment Monitoring Systems (PRAMS). Of course, they start with a sampling frame of birth certificates, whose mandatory elements can be regarded as a cultural relic of earlier times when telephone numbers were nonexistent or rare. Even so, the proportion of new mothers living in households without telephones cinches Shulman and Adams' choice of the mailed questionnaire as the better primary collection method.

This does not mean that the mailed questionnaire with telephone followup is necessarily the best data collection method nor even a wholly satisfactory one. The PRAMS data series was initiated to estimate the prevalence of various problems in pregnancy and early infancy; but even with their preferred collection method those classes of new mothers most likely to experience problems have the lowest response rates. Black women, women with less than a high school education, unmarried mothers, and mothers of babies of very low birth weight have reported response rates no higher than 70 percent in most States and under 50 percent in some. The classic race, education, and income biases of mail questionnaire surveys are still with us. The authors say nothing, at least in this paper, about how they plan to cope with the potential response biases which may result. Other solutions may be available, but one obvious answer is another followup round using personal interviews.

Since computer-assisted personal interviewing (CAPI) is a very new and rapidly developing technology, most papers on CAPI focus primarily or exclusively on operational rather than methodological questions. Couper, Groves, and Kosary attempt to broaden the discussion and introduce more sophisticated forms of methodological thinking into the topic. They also provide a useful review of the small literature on CAPI and a convenient summary of potential "issues" in CAPI development. "Issues" may be somewhat too strong a word, however, for the typically unresolved open questions in this fledgling field.

How new is CAPI? To my knowledge there is only one continuing national survey collecting data exclusively with CAPI, the Netherlands Labor Force Survey (van Bastelaer et al., 1988). In the U.S. two national surveys have used CAPI to collect part of their data: (1) the Nationwide Food Consumption Survey conducted by National Analysts for the U.S. Department of Agriculture (Rothschild and Wilson, 1988); and (2) the National Health Interview Survey conducted by the Census Bureau for the National Center for Health Statistics. Neither of these two U.S. surveys involve case management by CAPI or electronic transmission of data.

There is broad interest in CAPI. Many government agencies and private survey organizations in the U.S., Australia, Canada, France, Sweden, and the U.K. are developing plans to use CAPI in the relatively near future. Several have conducted small CAPI tests.² The idea is very appealing. For surveys currently employing both personal and telephone interviews, the use of CAPI to complement CATI is an obvious next step. As a concept, CAPI also has been winning rapid support among the upper management levels of survey organizations, perhaps because as Couper et al. suggest (at least in an early draft of their paper), CAPI's use of existing field staff suggests that the transition to CAPI will be less disruptive, and perhaps less expensive, than the transition to CATI. This, I believe, is a misconception.

There are many practical hurdles to implementation of CAPI (or an integrated CATI-CAPI system). Couper et al. have listed many of them. I would like to re-emphasize four: (1) lack of existing hardware meeting common field requirements for weight, battery life, handling ease,

and memory; (2) hardware volatility risking early obsolescence of any equipment procured; (3) software requirements for case management, supervision, and communications far exceeding those for CATI; and (4) uncertain effects on field operations, field supervision, and field organization. At present I believe we have something of a "CAPI-Gap" existing between management-supported aspirations for CAPI and the knowledge and resources necessary to design and efficiently implement functioning CAPI systems that interviewers, supervisors, researchers, and administrators can live with.

I disagree with the Kosary, Groves, and Couper paper only on two related points. First, I am not sure CAPI is ready for sophisticated experimental designs. Couper et al. suggest that the principal limiting factor in CAPI research is the expense of such studies. At this point, I suspect that operational problems are a greater obstacle. It is sufficiently difficult to assemble the necessary hardware, software, and training materials to conduct a CAPI survey reliably with one set procedures. Concurrent preparation for two or more kinds of hardware, software, or procedures in experimental designs may be beyond the capabilities of most organizations at this time. With more experience, of course, such designs will be possible.

Second, I wonder how useful sophisticated methodological testing is for CAPI at the present time when its basic operating procedures are uncertain and ill defined. As in most new areas of research, it may be better to begin with rough observational methods and loose field trials until more practical experience is gained. More sophisticated forms of methodological testing may not be appropriate until operational procedures are under better control.

To illustrate the value of experience, (and to provide some taste of actual problems) I would like to conclude by describing Census Bureau experiences with CAPI in the 15-minute AIDS supplement which was conducted as part of the 90-minute predominantly paper-and-pencil Health Interview Survey. Each interviewer was equipped with a Gridlite laptop for the AIDS supplement but each also was provided with paper forms for use when necessary, such as in high crime areas.

After the interviewers were fully trained and submitting AIDS supplements on diskettes on a regular basis, Neil Ferraiuolo of the Census Bureau staff discovered that in some Census Bureau regions, the interviewers were completing up to 50 percent of the AIDS supplements on paper forms and then keying them into their laptops at the end of the day. This met the client agency's goal of transmitting machine-readable data from the field, but not in the way planned. What was intended as a computer assisted interviewing system had become in many cases a home key entry system.

None of our advance speculations nor two prior tests of CAPI prepared us for this outcome, perhaps because the interviewers in prior tests frequently were accompanied by observers from headquarters or from client agencies. This was our first use of CAPI where the interviewers were left largely on their own under their regular supervisors. The Hawthorne effect was removed. The specialized use of CAPI in this survey,

limited to one supplement of a predominantly paper-and-pencil interview, also may have contributed to the unexpected interviewer reaction.

Reasons the interviewers gave for not using their laptops for the supplement were often surprising either in their content or in their frequency. The most common explanations, for example, referred to the weather -- it was too hot, too cold, rainy, or threatening to rain. The interviewers reportedly feared the laptop would be damaged and they could be charged with its replacement cost. They protected this valuable piece of government property by leaving it safely at home.

Health reasons also were common and are difficult to circumvent when raised. The interviewer had a bad back, a sore arm, trouble seeing, or her doctor told her not to carry heavy objects. The two minutes required to boot up the laptop and start the supplement also were seen as a threat to respondent cooperation, so paper forms often were used when respondent cooperation was uncertain. Paper forms also were preferred in telephone followups, especially when (as in most motel rooms) the telephone wasn't located on a table large enough for a laptop computer. A number of additional practical problems were mentioned, some anticipated and some not.

This experience suggests both a host of operational problems to be solved and a behavioral metric of operational success. Until our interviewers voluntarily choose CAPI for virtually all their interviews, the operational problems of this new technology require more work. Couper, Groves, and Kosary have proposed laboratory testing and surveys of interviewers and respondents as methods to learn more about CAPI and its effects. These methods may help; but major advances at this stage, I suspect, will occur primarily through repeated experience in the field. The more practical experience we gain, the more valuable our think pieces on CAPI design and its methodological effects will become.

Footnotes

¹The views expressed are those of the author and do not necessarily reflect official views or policies of the U.S. Census Bureau.

²In addition to the studies cited above and by Couper et al. cf.: Bernard (1988); Danielsson and Maarstad (1982); Nicholls (1988); and Rice et al. (1988).

References

Bernard, Catherine (1988), "Two, New, Experimental Collection Methods at INSEE," in J. Kovar and E. Doucet, Proceedings of Symposium 88: The Impact of High Technology on Survey Taking, Ottawa, Canada: Statistics Canada.

Danielsson, L, and P.A. Maarstad (1982), "Statistical Data Collection with Handheld Computers - A Test in Computer Price Index," unpub-

lished report of Statistics Sweden, Orebro, Sweden.

Groves, Robert M. and Robert L. Kahn (1979), Surveys by Telephone, New York: Academic Press.

Hubble, David L. and Bruce E. Wilder (1988), "Preliminary Results from the National Crime Survey (NCS) CATI Experiment," Proceedings of the Survey Research Methods Section, American Statistical Association, pp. 558-563.

Lepkowski, James M. (1988), "Telephone Sampling Methods in the United States," in Robert M. Groves et al., Telephone Survey Methodology, New York: Wiley, pp. 73-98.

Nicholls II, William L. (1988), "Discussion" of papers by Rothschild and Wilson and by Sebestik et al., Proceedings of the Bureau of the Census Fourth Annual Research Conference, Washington, U.S. Department of Commerce, Bureau of the Census, pp. 340-342.

Rice Jr., Stewart C., Robert A. Wright, and Ben Rowe (1988), "Development of Computer Assisted Personal Interview For the National Health Interview Survey," Proceedings of the Survey Research Methods Section, American Statistical Association, pp. 397-400.

Rothschild, Beth B. and Lucy B. Wilson (1988), "Nationwide Food Consumption Survey 1987: A Landmark Personal Interview Survey Using Laptop Computers," Proceedings of the Bureau of the Census Fourth Annual Research Conference, Washington, U.S. Department of Commerce, Bureau of the Census pp. 347-356.

Schwanz, Dennis, Ed Montfort, and John Cannon (1988), "Analysis of Operational Issues: 1987 AHS CATI," CATI Research Report No. AHS-1, U.S. Bureau of the Census, Computer-Assisted Interviewing Central Planning Committee, CATI Research and Analysis Subcommittee.

U.S. Census Bureau (1988A), Final Report on Planning for a Post-1990 Redesign of the Current Population Survey. A report of the Oversight Committee of the Joint BLS-Census Committee on The Current Population Survey.

U.S. Census Bureau (1988B), Evaluation of CATI Data Quality and Costs in the Current Population Survey, CATI Research Report No. CPS-2 of the Computer-Assisted Interviewing Central Planning Committee, CATI Research and Analysis Subcommittee.

U.S. Office of Management and Budget (1984), "The Role of Telephone Data Collection in Federal Statistics," Statistical Policy Working Paper 12.

van Bastelaer, Alois, Frans Kessemakers, and Dirk Sikkel (1988), "Data Collection with Hand-Held Computers: Contributions to Questionnaire Design," Journal of Official Statistics, Vol.4, No. 2, pp. 141-154.

Waksberg, Joseph (1978), "Sampling Methods for Random Digit Dialing," Journal of the American Statistical Association, Vol. 73, No. 361, pp. 40-46.