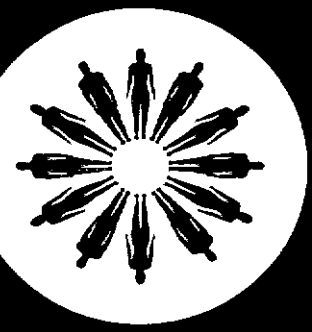


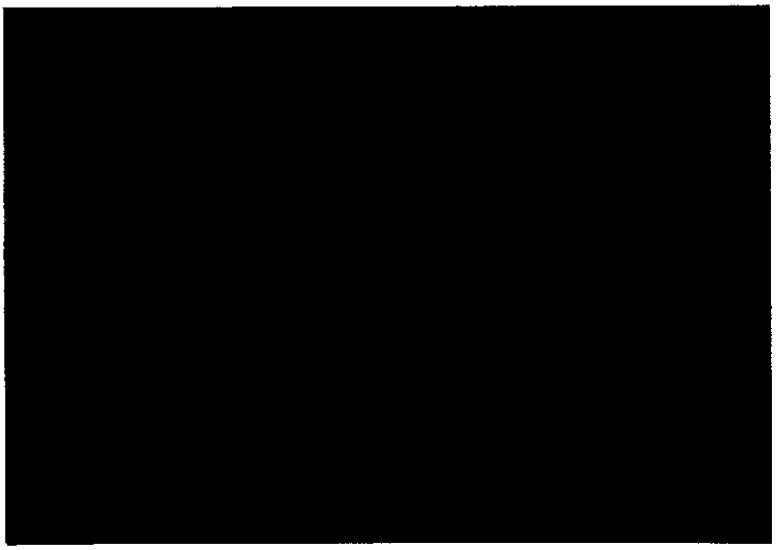
National Longitudinal Surveys

U.S. Department of Labor
Bureau of Labor Statistics

Discussion Paper

The National Longitudinal Surveys (NLS) program supports many studies designed to increase understanding of the labor market and methods of data collection. The *Discussion Papers* series incorporates some of the research coming out of the program. Many of the papers in this series have been delivered as final reports commissioned as part of an extramural grant program, under which grants are awarded through a competitive process. The series is intended to circulate the findings to interested readers within and outside the Bureau of Labor Statistics.

Persons interested in obtaining a copy of any *Discussion Paper*, other NLS reports, or general information about the NLS program and surveys should contact the Bureau of Labor Statistics, Office of Economic Research, Washington, DC 20212-0001 (202-606-7405).



NLS Discussion Papers

Two Papers on the Use of Computer-Assisted Personal Interviews in the National Longitudinal Survey of Youth

N. M. Bradburn
M. Wojcik
A. Schoua-Glusberg
M. R. Pergamit
M. Frankel
J. Ingels
E. Hunt
R. P. Baker

May 1991

Report: NLS 92-2

A Comparison of Computer-Assisted Personal Interviews (CAPI) with Paper-and-Pencil (PAPI) Interviews in the National Longitudinal Study of Youth

Norman M. Bradburn, Martin R. Frankel, and Reginald P. Baker
National Opinion Research Center
University of Chicago

Michael R. Pergamit
Bureau of Labor Statistics

May 1991

This paper was prepared for presentation at the annual meetings of the American Association for Public Opinion Research held in Phoenix, Arizona, May 16-19, 1991. Opinions stated in this paper do not necessarily represent the official position or policy of the U.S. Department of Labor.

NORC

A Comparison of Computer-Assisted Personal Interviews (CAPI) with Paper-and-Pencil (PAPI) Interviews in the National Longitudinal Study of Youth

Norman M. Bradburn, Martin R. Frankel, Reginald P. Baker

National Opinion Research Center

University of Chicago

Michael R. Pergamit

Bureau of Labor Statistics

AAPOR Conference

Phoenix, Arizona

May 16-19, 1991

In discussions of mode effects, the survey methodology literature distinguishes three modes of data collection--face-to-face, telephone and self-administered. There is an extensive literature on possible effects of collecting data by each of these modes because they appear to differ in fundamental ways. What has been less noticed, however, is that there are variations within each of these methods regarding whether or not they are computer-assisted; that is, whether the questionnaire is represented in electronic or paper-and-pencil form. There is a paucity of literature on within-mode effects of using computers to assist in the data collection process.

There are several reasons for this neglect of research on the effects of computer assistance. Until recently there has been relatively little experience with computer-assisted data collection outside the telephone mode. While the possibility of using computers to assist face-to-face interviewing has been recognized for a long time, its reality has awaited the development of light, portable, inexpensive computers that ordinary interviewers could carry around with them and use with relatively little training. Only in the last few years have such computers been available and has CAPI become a practical reality. Computer-assisted self-administered questionnaires have been used in special settings such as in mall intercept studies or in medical offices to take medical histories. As with CAPI, computer-assisted self-administration (CASAQ) has been hampered by the lack of inexpensive portable computers and widespread home ownership of compatible computers that could be relied upon for use with questionnaires sent to individuals on floppy disks. Thus, in effect, studies of differences between CAPI and PAPI have been limited to the telephone mode.

But computer-assisted telephone interviewing (CATI) began almost simultaneously with the spread of telephone interviewing, and indeed in some people's minds is synonymous with telephone interviewing. Shifting from face-to-face to telephone interviewing was viewed as such a big difference because of the changes in the way respondents and interviewers interacted and the differences in the availability of visual cues in the two modes. Other possible differences, such as whether the questionnaire was represented in electronic form, were not viewed as serious sources of response variation. There was not a long history of

doing studies on the telephone, so that changes from PAPI to CATI were not seriously investigated.

With CAPI, however, the situation is quite different. We have decades of experience with face-to-face PAPI and are only now beginning to explore the feasibility of CAPI. If CAPI is introduced in a survey that has been done for a number of years with PAPI, as is the case that I am reporting on today, it is extremely important to know whether observed changes are due to mode effects or to real change in the phenomena. I hope that the spread of CAPI will be accompanied by a vigorous research program designed to investigate mode effects. If there turn out to be effects that are associated with the introduction of the electronic representation of the questionnaire, then perhaps we should look harder at whether or not something similar happens within telephone interviewing.

In discussing mode effects, we can distinguish among three types of effects--those that change the interviewer's behavior, those that change the respondent's behavior, and those that change in the interaction between the interviewer and the respondent. The most obvious effects of CAPI are those that change the interviewer's behavior because it is the interviewer that is most affected by a change from PAPI to CAPI. Indeed, it is not immediately obvious that there should be any effect on respondents' behavior because, from their point of view, they are getting the same questionnaire as they would if the interviewer were working with a paper-and-pencil representation. The use of a computer for recording answers, however, may change the way respondents view the task and thus have an impact on their behavior. Finally, reading questions off a computer screen and typing in responses may change the quality of the interaction between interviewer and respondent, for example by reduced eye contact or an increased formality in which the computer becomes a third party to the interview.

THE ROUND 11 EXPERIMENT

The NLS/Y is a longitudinal face-to-face survey administered by the Bureau of Labor Statistics (BLS) now in its 13th year. The sample consists of 11,464 people who were aged

14-21 as of January 1,1979 and who have been interviewed every year since then. The sample was stratified by sex, race, ethnicity, and poverty status with oversampling of blacks, hispanics, white youths in poverty and equal numbers of men and women in each group. The questionnaire is primarily oriented toward labor market participation, education and fertility. It is an extremely complex questionnaire with many skip patterns and extensive rostering. The complexity of the questionnaire and the difficulties that interviewers have in making their way through it correctly make it an ideal candidate for a computer-assisted form.

NORC began work on a CAPI version of NLS in 1989 using software developed by the Center for Human Resources Research at Ohio State, the prime contractor for this study. In our first experiment, half of the Ohio sample for Round 11 of NLS, randomly selected from all Ohio cases, was interviewed using CAPI. In all, 301 cases were completed by CAPI (completion rate = 91.8%) and 264 cases were completed by PAPI (completion rate= 95.6%). Because some of the cases were done on the telephone, the data presented here is only for those cases in the experiment that were completed face-to-face (260 CAPI and 247 PAPI).

All 20 Ohio interviewers were trained on CAPI. They completed approximately half of their cases using CAPI (roughly 16 CAPI cases per interviewer) and half using PAPI. The PAPI cases were supposed to be done first so that the interviewers were already trained on the instrument and had experience with Round 11 cases. Training on CAPI was conducted after the interviewers had already done some PAPI cases.

A second larger nationwide experiment was conducted as part of Round 12. Data from that experiment will be reported in subsequent papers.

RESULTS

The NLS/Y is an extremely complex questionnaire, and there are a vast number of comparisons that can be made between the PAPI and CAPI modes. In this paper I shall examine hypotheses drawn from other mode effects studies to see if some of the same types of effects occur here.

First, however, I can report that out of 139 variables we examined, we only found 4 that reached conventional levels of significance. Of these 4, two are related to reports of alcohol use, one concerned the unit of pay and one was respondent attitude toward the interview, which is not part of the NLS questionnaire itself. I shall organize the discussion around 5 types of possible effects--interviewer error, length, anonymity, respondent motivation and the venerable "other".

Interviewer Error

The major advantage of computer representation of questionnaires is that the investigator can better control the presentation of questions and subsequent branching patterns that are contingent on the responses. In complex questionnaires like NLS/Y there are many skip patterns and sections of the questionnaire that refer to responses earlier in the questionnaire. Even with extensive training, interviewers make many mistakes as they proceed with the interview. Many of the these mistakes are caught by the interviewer during the interview itself, others are corrected in the central office by editors, but some slip through and cannot be corrected even with extensive cleaning. One of the primary attractions of CAPI is the ability to program skips so that branching errors are eliminated. Errors of this type exist only insofar as the programming has not been able to anticipate every possible combination of answers and allowed for a proper track through the questionnaire. Reducing this type of error can improve the quality of the data and reduce costs by eliminating costly editing and retrieval and reducing the extent of cleaning.

CAPI does have a big effect in this domain. Table 1 presents the average error rates for the CAPI and PAPI cases. It is clear that CAPI is working the way it is supposed to and reducing the number of interviewer skip errors.

TABLE 1 HERE

Note that interviews done with CAPI also had fewer legitimate missing "Don't Knows" and item refusals. It is not clear why this difference should occur. Perhaps respondents view CAPI as a more anonymous form of data recording and are more willing to respond to some questions. This hypothesis would be consistent with some suggestive differences described below. We will have to follow up in our analyses of Round 12 to see whether this finding is replicated and whether we can understand better what might cause it.

Length

One of the major differences between face-to-face interviewing and telephone interviewing is length. Interviews conducted over the phone tend to go faster and responses to individual questions tend to be shorter. Indeed, this difference is one of the primary mode effects. Does anything similar happen with CAPI? The answer is clear and simple--No. Table 2 shows the average length of interview for the two modes.

TABLE 2 HERE

CAPI cases averaged 50.2 minutes/case; PAPI cases averaged 51.1 minutes/case with large standard deviations for both groups. We do not have timing for individual items, as one can from telephone interviewing, but times for some sections in Round 12 were recorded, and preliminary analysis suggests that they also did not show any difference.

Anonymity

One of the most interesting mode differences among face-to-face, self-administered and telephone interviewing found by Sudman and Bradburn (1974) in their review of response effects was on sensitive questions where there may be problems of self-presentation. With more private modes of administration, e.g. self-administration, respondents were more likely to report negative behavior. While it is not apparent that CAPI would be viewed as more anonymous than PAPI, since both are face-to-face, it is possible that the fact that the interviewer is entering the responses into a computer rather than writing them down on a form that has the respondent's identifying information on it, will make respondents feel more

anonymous and encourage them to report more negative information. There is some evidence that this may be the case, as is shown in TABLE 3.

TABLE 3 HERE

Table 3 presents the frequencies for the CAPI and PAPI cases on a number of items relating to alcohol use, questions that have been shown to be sensitive and affected by wording and format changes. (see Bradburn, Sudman and Associates, 1979.) While not all of the items show differences in the direction of higher frequencies for CAPI and most the differences are probably not statistically significant (2 of them are), in 23 out of 26 comparisons reports of frequency of alcohol-related problems in the last year are greater for CAPI than for PAPI cases. The findings suggest a pattern that is worthy of further investigation. Data from Round 12, which has a larger sample size, will enable us to follow up on this possible effect.

Motivation of respondents

The quality of data is affected by respondents' motivation to provide the requested data, particularly in a study like NLS/Y which taxes the respondents' memories for reports of their past behavior in many different areas and asks a number of sensitive questions, not only about such things as alcohol use but also about detailed information on income and assets. If the mode of administration of the survey affects respondent motivation, it can thus have a profound impact of the quality of data. Insofar as we can detect any effects of CAPI on respondent motivation, it would appear to be positive or neutral.

TABLE 4 HERE

There was no difference in the response rates between the two modes, a finding that is replicated in the Round 12 experiment. Since these respondents have been interviewed for so many years and have continued to cooperate, response rate is not a very sensitive measure of motivation. There was, however, a difference in the interviewers' ratings of respondents'

attitude toward the interview with the CAPI respondents being rated as more friendly and interested than the PAPI respondents. In Round 12 we obtained more data from respondents about their reaction to CAPI and may be able to say more about the motivational effects of CAPI.

Other effects

I shall end with a puzzle. One unexpected difference that has emerged from the examination of the data concerns the proportion of respondents who report being paid by the hour. As can be seen in Table 5, 76% of the CAPI respondents report being paid by the hour while only 67% so report when the interviews were done PAPI.

TABLE 5 HERE

This difference in wage rates is not statistically significant, but the proportion being paid by the hour is probably significant. Since we have examined a large number differences, and we would expect some significant differences on the basis of chance, it is possible that this finding is just one of those chance differences. But there are indications from some preliminary analysis of the Round 12 experiment that this difference is appearing there as well. So it looks as if it is something real.

If it is a real difference, it is not clear what is causing it. The questions asked of respondents about this topic are complex and have several subparts to enable respondents to answer in a flexible fashion.

FIGURE 1 HERE

In the PAPI version all of the questions appear on one page together so that the interviewers can see what is going to be asked and where the question is going. CAPI interviewers, on the other hand, get only one question at a time on their screens, with the next question being determined by the response to the first question. Since the set of questions

involve income, pay rates and whether one is paid by the hour or some other time period, respondents may answer the questions before they are asked. The PAPI interviewers, knowing what is coming, might actually use the suggested probe more often and get more hourly reports in the beginning. But if that were the case, there would be a higher proportion reporting being paid by the hour when the question is first asked. As we seen Table 5, however, the difference shows up only in response to the second question in which they were asked specifically about being paid by the hour. It is possible that interviewers in the PAPI condition were not in fact asking the follow-up question, but were coding the answer as "no" on the basis of either what they had heard before the question would have been asked or their assumption that respondents who reported incomes as paid by the week or month were not paid on an hourly basis. Further analysis is needed before we can understand what is happening with this question. For Round 12 we have tapes of both CAPI and PAPI interviews, so we will be able to test out some of these process hypotheses.

Conclusion

We examined the differences for 139 variables between CAPI and PAPI cases in an experiment where assignments had been made randomly to mode of administration. Except for effects on interviewer errors that were programmed into the CAPI itself, in all of these comparisons we found only 4 differences that looked as if they might even approach statistical significance. This number is within the number that one might expect by chance when making multiple comparisons.

There are a few differences, however, that deserve further study before rejecting them. One intriguing hint is that respondents may treat CAPI as a more anonymous mode of data collection and increase their reports of sensitive or negative behavior. A second possible finding is that there may be some question formats designed to give flexibility to respondents that when available to the interviewer only one question at a time affects the way respondents answer or how their answers are recorded. Considerable further work is needed to understand what is really going on in this case.

NORC

Table 1

Missing Data

	PAPI (N=260)	CAPI (N=248)
Invalid skips	.19%	0%
Don't know	.43%	.15%
Refusals	.19%	.004%

NORC

Table 2

Length of Interview

	Average Minutes/ Case	S.D.	N.
PAPI	51.1	27.4	247
CAPI	50.2	17.5	259
TOTAL	50.7	22.9	506

NORC

Table 3

Alcohol-Related Behavior Reports

Question	PAPI (N=161)	CAPI (N=149)
21/60 # times had 6 or more drinks in last 30 days 4 or more	17	26
21/65 Ever had strong desire or urge to drink Yes	37	42
21/66 Ever drink much more than intended 1 or more times during the year	42	55
21/67 Ever found it difficult to stop drinking Yes	12	19
21/68 Ever driven car after drinking too much Yes	22	31
21/69 Ever been sick the morning after drinking 1 or more times during the year	21	30
21/70 Drinking caused R or someone to be hurt 1 or more times during the year	14	12
21/71 Ever drink for a long period of time 1 or more times during the year	19	33

NORC

Table 3

Alcohol-Related Behavior Reports (Con't.)

Question	PAPI	CAPI
21/72 Need more to get drunk now than before 1 or more times during the year	14	20
21/73 Spouse/ partner threatened to leave because of drinking 1 or more times during the year	4	5
21/74 Tried to cut down drinking but couldn't 1 or more times during the year	7	8
21/75 Ever shake the morning after drinking 1 or more times during the year	8	7
21/76 Given up on activities/ interests to drink 1 or more times during the year	4	5
21/77 Needed drink so badly couldn't think 1 or more times during the year	3	3
22/10 Drink more than before to get same effects 1 or more times during the year	7	11
22/11 Stayed away from work because of drinking 1 or more times during the year	3	9

NORC

Table 3

Alcohol-Related Behavior Reports (Con't.)

Question	PAPI	CAPI
22/12 Lost ties with family because of drinking 1 or more times during the year	1	5
22/13 Drinking instead of doing something you were supposed to 1 or more times during the year	11	18
22/14 Kept drinking even though threat to health 1 or more times during the year	6	11
22/15 Chance for raises hurt by R's drinking 1 or more times during the year	0	1
22/16 Spend a lot ot time drinking 1 or more times during the year	6	14
22/17 Hangover interfered with something else 1 or more times during the year	9	15
22/18 Kept drinking after emotional problem 1 or more times during the year	3	9
22/19 Heard/seen things not real after drinking 1 or more times during the year	1	5

NORC

Table 3

Alcohol-Related Behavior Reports (Con't.)

Question	PAPI	CAPI
22/20 Taken a drink to keep from shaking 1 or more times during the year	1	5
22/21 Kept drinking after problems at home 1 or more times during the year	4	5

NORC

Table 4

Respondent Reactions

	PAPI (N=247)	CAPI (N=260)	TOTAL (N=507)
Response rate	95.6%	91.8%	93.59%
Respondent attitude toward interview			
Friendly/ interested	80	91	85
Cooperative	18	7	12
Impatient	1	2	3
Hostile	1	0	--
	100	100	100

NORC

Table 5

Respondent Pay Unit

	PAPI (N=137)	CAPI (N=172)	TOTAL (N=309)
Identified as hourly in first question	49%	48%	49%
Identified as hourly in second question	18%	28%	23%
Total hourly	67%	76%	72%
Average hourly wage			
When identified in the first question	$6.81	$6.82	$6.82
When identified in the second question	$8.03	$7.92	$7.96
Total	$7.12	$7.22	$7.17

NORC

FIGURE 1

A. **Altogether, how much (do/did) you usually earn at that job? ENTER IN APPROPRIATE BOXES. PROBE IF NECESSARY: Was that per hour, per day, per week, or what?**

/___/___/___/ , /___/___/___/ . /___/___/
DOLLARS CENTS

Per hour	(GO TO Q.23)	01
Per day	(GO TO B)	02
Per week	(GO TO B)	03
Bi-Weekly (Every 2 weeks)	(GO TO B)	04
Per month	(GO TO B)	05
Per year	(GO TO B)	06
Other (SPECIFY) _____	(GO TO B)	07

B. **(Are/Were) you paid by the hour on this job?**

Yes	(ASK C)	1
No	(GO TO Q. 23)	0

C. **How much (do/did) you earn per hour?**

/___/___/___/ . /___/___/
DOLLARS CENTS

A COMPARISON OF COMPUTER-ASSISTED PERSONAL INTERVIEWS (CAPI) WITH PERSONAL INTERVIEWS IN THE NATIONAL LONGITUDINAL SURVEY OF LABOR MARKET BEHAVIOR-YOUTH COHORT

N.M. Bradburn, M. Frankel, E. Hunt, J.Ingels,
A. Schoua-Glusberg, M. Wojcik
National Opinion Research Center
and
Michael R. Pergamit
Bureau of Labor Statistics

ABSTRACT

The purpose of this experiment was to assess the effect of conducting interviews in Round 12 of the NLS-Y by the Computer-Assisted Personal Interview (CAPI) method as compared with the traditional paper-and-pencil personal interview method. The experiment was conducted on one-half of the total sample and excluded respondents who had to be interviewed outside the United States and/or in Spanish. Interviewers were assigned cases in the same geographical region and, where possible, were matched with respondents for ethnicity. Assignment to the proper experimental or control group was done through random assignment of interviewers. Thus the experiment reflects actual field practices. The paper will report on the operational problems in conducting the experiment.

KEYWORDS

CAPI; Mode Effects

INTRODUCTION

In this paper we present the first report of an experiment conducted in Round 12 of the National Longitudinal Survey of Labor Market Behavior-Youth Cohort(NLS) in which a computer-assisted personal interview method (CAPI) was compared with the standard method of face-to-face paper-and-pencil personal interviewing (PAPI). The purpose of this experiment is to assess the effect of conducting interviews by the CAPI method as compared with the PAPI method. Previous pilot work has demonstrated the logistical feasibility of conducting NLS interviews with CAPI, but no systematic studies have been done to investigate possible method effects on interviewer accuracy or respondents' responses.

THE NLS AND PREVIOUS CAPI EXPERIENCES

The NLS is a longitudinal face-to-face survey administered by the Bureau of Labor Statistics(BLS) now in its 13th year. The Round 12 sample consists of 11,465 people who were aged 14-21 as of January 1,1979 and who have been interviewed every year since then. The sample was stratified by sex, race, ethnicity and poverty status with oversampling of blacks, hispanics, white youths in poverty and equal numbers of men and women in each group. The questionnaire is primarily oriented toward labor market participation, education and fertility. It is an extremely complex questionnaire with many

skip patterns and extensive rostering. The complexity of the questionnaire and the difficulties that interviewers have in making their way through it correctly make it an ideal candidate for a computer-assisted form.

NORC began work on a CAPI version of NLS in 1989 using software developed by the Center for Human Resources Research (CHRR) at Ohio State, the prime contractor for this study. In a pilot study, half of the Ohio sample for Round 11 of NLS, randomly selected from all Ohio cases, was interviewed using CAPI. In all, 301 cases were completed by CAPI (completion rate = 91.7%). All 20 Ohio interviewers were trained on CAPI. They completed approximately half of their cases using CAPI (roughly 16 CAPI cases per interviewer) and half using PAPI.

While there were some problems that developed during this pilot study, the pilot was considered successful and a more ambitious experiment was conducted in Round 12. In the pilot study we used the Toshiba 1200 Laptop Computer (LTC) with a 20MB hard disk drive, a floppy disk drive, back-lit screen and 2400 baud internal modem weighing 14 lbs. which could be powered by battery or electrical outlet. From the pilot study we learned that:

* At 14 lbs. LTC weight is a problem but is not an obstacle to use of the computers.

* Use of battery vs electric power presented no problems to either the interviewers or respondents.

* Hardware problems were reported by about 5% of the interviewers, mostly during the first two weeks of the field period.

* Screen readability presented a problem, particularly in bright light.

Interviewer difficulty with the keyboard turned out to be due to lack of typing skill, rather than any special problem with computer keyboards. Practice improved performance.

Software

The software used in these experiments was especially developed for this study by the CHRR staff, under the leadership of Randall J. Olsen, rather than adapting a previously used computer-assistance programming system as has been common in previously reported CAPI experiments(Denteneer, et al.,1987; Rothschild and Wilson, 1988; Sebestik et al, 1988). The software incorporated the basic CADAC features of controlling skip patterns and ranges, rostering, allowing movement within the instrument, "filling" preentered data, and other minor features. It also handled the complex calendar operations in the employment history section--a critical section which had given interviewers considerable difficulty in the PAPI mode.

It is, of course, an advantage to be able to develop CAPI software for a specific study, but it still does not make it an easy task. The development and testing of software requires considerable effort and extensive cooperation between computer programmers and survey specialists. It cannot be viewed as

simply designing a questionnaire appropriate for PAPI and then turning it over to the programmers.

Software development began with NORC survey staff advising the programming staff on the requirements of the software for actual field situations. When CHRR programmers completed internal testing and judged a software version ready for further testing, the NORC staff began testing by "keybanging" to determine if the software could be induced to break down. In addition the software was tested for screen formats, consistency between PAPI question text and CAPI question text, acceptability of ranges, enforcement of consistency between question responses, correctness of skip patterns, correctness of rostering and awkwardness of usage. When faults were found in the software, programming staff first attempted to duplicate the error to verify its existence and then correct it.

Most problems were found during the testing stage and quickly corrected. A few more were discovered by the interviewers during training and were corrected before the field period began. During the first 2 weeks of the field period several other problems were discovered and corrected by instructing the interviewers to phone in electronically to receive "fixes". The "fix" was made by the CHRR programmer and sent to NORC for checking to verify its accuracy and that it did not in turn create additional errors. When the fix was successful, the CHRR programmer set up a program so that the next time an interviewer dialed in, the fix was automatically put into her computer.

While one would like to have the CAPI program completely correct by the time the field period begins, it is likely that some bugs will show up during actual field interviewing even with the most rigorous testing beforehand. Pencil-and-paper questionnaires frequently have errors also. The possibility of fixing them after the field period begins is small. With CAPI, however, updating of the programming was possible during the field period and was done electronically by the central office without having to rely on interviewers' correcting the programs in their own LTC's.

THE NLS-12 EXPERIMENT

The success of the Round 11 pilot led BLS to consider usage of CAPI for the NLS. They felt, however, that it was first necessary to acquire more information: (1) What steps were necessary to carry out CAPI on a national scale?; (2) What was the impact on the costs of carrying out the survey?; and (3) What was the impact on the data of a change in collection method?

Due to cost considerations, it was necessary to carry out further CAPI testing in a live-production framework. BLS determined that a controlled experiment was required. Information would be collected which would allow us to evaluate each of the three questions posed above.

The NLS-12 experiment was designed with two constraints in mind. First, it was to be a true experiment and strict procedures were developed to insure that there was no contamination between the experimental and control groups; and

second, the experiment could not compromise the quality of the data for the study as a whole, including the overall completion rate target of 92% which was necessary to maintain the integrity of the panel.

The experimental design called for division of the NLS sample into three subsamples: 25% of the respondents were interviewed using CAPI, 25% were designated as a control group to be interviewed using PAPI; the remaining 50% were regular NLS cases that were not part of the experiment. The data from the CAPI and Control cases will be compared in the analyses of mode effects. The design serves as a substitute for overlapping samples which one might use in a repeated cross-section survey. Thus, any mode effects from CAPI which cannot be corrected can be controlled for in any future use of the data.

Because random selection of cases into the two groups would not result in cost efficient clustering of cases for interviewer assignment, mode selection was based on the attributes of the interviewers and their assignment characteristics e.g. geographical location and previous NLS experience. Thus, 25% (77) of the Round 12 interviewers were randomly assigned to the CAPI condition, 25% (77) were assigned to the control condition and the remaining 170 interviewers were assigned to the regular cases which were also done PAPI.

The sample selection process was as follows: field managers assigned NLS cases to each interviewer as in any normal round. Interviewers were characterized by assignment size and complexity (the presence or absence of supplementary child cases), urban vs rural case load, assignment type (regular vs regular-plus-conversion cases), and ethnicity (black, white, Hispanic). Interviewers with Spanish language cases and cases outside the 48 contiguous states were excluded from the experiment for cost reasons, the anticipated difficulty with international phone transmission of cases, and the cost of providing interviewer support by international long distance. These types of cases are a very small portion of the NLS sample and excluding them from the experiment was judged to have little adverse impact on the generalizability of the experimental findings.

Interviewers were randomly assigned to conditions by using procedures which balanced the conditions on the basis of interviewer type: regular interviewer vs converter; ethnicity; geographic region; and metropolitan/non-metropolitan areas. This was accomplished by sorting the interviewers with respect to these four variables and then forming groups of four "matched" interviewers. Within each matched group, interviewers were randomly distributed among 4 groups: 1 CAPI group, 1 control group, and 2 regular groups.

Cases belonging to a CAPI condition interviewer were done in the CAPI mode, those belonging to a control interviewer were in the control condition and done PAPI. The remaining cases were "other" and done PAPI. To maintain experimental integrity interviewers and cases could not cross mode, that is a CAPI interviewer could not interview a respondent in any other condition as long as the interviewer was participating in the experiment. No PAPI interviewers were allowed to do cases using CAPI.

The total number of cases assigned to the CAPI mode was 2814; the control sample had 2715 cases.

CONTROL OF THE SAMPLE DESIGN

Because interviewers were not allowed to work in more than one mode, certain of the usual field economies were not possible in the Round 12 data collection effort. For instance it was not possible to use one interviewer to conduct interviews with two neighboring respondents if the respondents were in different modes. The sample selection process (selecting assignments rather than cases) minimized these effects. Further because of the exclusive nature of the experimental modes and the use of a computer supported field management system, it was possible to track experimental costs efficiently and easily.

The NORC automated Field Management System (FMS) is designed to track, on a case by case basis, production, cost and sample information. Modifications in the system set up for Round 12 were designed to allow tracking of the experiment and to assure that no contamination occurred between experimental conditions.

Executing a field experiment of this magnitude requires continual tracking and the understanding and cooperation of the field management staff. Initially there was some resistance on the part of field managers to some of the requirements of the experiment since it meant that cases could not be freely transferred from interviewer to interviewer to maximize completion rate and minimize costs. Special attention had to be given to explaining to the field supervisors the logic of the experimental method and the necessity of keeping experimental and control cases completely separate. A special procedure was established so that each week requests for reassignment of cases between interviewers were reviewed by the principal investigators to make sure that no reassignments were made that would compromise the experiment.

Near the end of the field period it became clear that, within the constraints of the budget, we could not achieve the targeted completion rate for the entire study and still maintain strict adherence to the experimental conditions. Thus, for cost reasons, we stopped the experiment after we had reached a completion rate of at least 80% each for the CAPI and the control conditions (2305 cases, or 81.9%, for CAPI and 2268, or 83.5%, for the controls). The remaining cases were completed by whatever method and interviewer was convenient. The overall completion rate for all Round 12 cases was 91.6%.

DATA FOR EVALUATING THE EXPERIMENT

A great variety of data were gathered as part of the experiment which will enable us to analyze potential mode effects. Interviewers filled out Interviewer Feedback Forms on each case. Problem Report Forms were used when problems were encountered as well as a lengthy 15th Case Questionnaire and an Exit Questionnaire. In addition, interviewers reported any problems to a CAPI Support Supervisor who assisted them in problem solving. In turn, the CAPI

Support staff kept the central office personnel apprised of all difficulties on at least a weekly basis. Respondents were also requested to fill out a very brief questionnaire on the CAPI experience.

Because the aim of the experimental evaluation of CAPI compared with PAPI data was to determine what--if any--mode effects occurred, the data had to be treated in a comparable manner once they reached the data reduction center. Once cases were received, the experimental design dictated somewhat different processing procedures to assure comparability of data. Control cases were receipted, edited, coded and data entered. A case selected for validation was sent to the validation shop; if it required retrieval, it went to a retrieval shop before being returned to the NLS library.

CAPI cases were electronically transmitted to NORC and immediately checked for problems; transmitted data were checked against data on the mailed-in diskette for discrepancies and a dump of verbatim information (occupation, industry, colleges attended, etc.) was printed out for coding. Dumps of validation questions were printed for cases selected for validation. It was anticipated that retrieval would be extremely rare and indeed this proved to be the case--there was only one CAPI case requiring retrieval. CAPI cases, then were not subjected to retrieval, editing or cleaning. Control cases were processed like regular PAPI cases except that after editing, but before cleaning, a first-stage file of control cases was created for the PAPI/CAPI evaluation. These data will be the core of much of the analysis of CAPI and PAPI edited cases.

Because we are interested in knowing whether CAPI affects the way interviewers read the questions or other aspects of the interactions between respondent and interviewer, tape recordings were made of the first, second and tenth interviews conducted by interviewers in the CAPI and Control conditions. These tapes will be coded for interviewer reading accuracy and for any mistakes in data entry by CAPI interviewers that would not produce a machine detectable error.

Data collection for the experiment ended in December and data processing is still in progress, so we are not able to report here on modal effects on the data themselves. We can, however, report on many operational aspects of conducting a large scale CAPI effort on a nation-wide basis.

OPERATIONAL ASPECTS

Hardware

NORC devoted a considerable amount of time and effort to the selection of the hardware that most closely met its needs as determined by the NLS-11 experimental results. From the interviewer's perspective, the most important desiderata were lighter weight computers and more readable screens. Ten computers were evaluated. The Compaq LTE was selected because of its light weight, backlit screen, internal modem and 20MB hard disk, and long battery life. Although the Compaq computers were priced slightly higher than some of

the others considered, they were selected because of their ability to meet all of our needs as well as the support offered by the vendor. Each CAPI interviewer was provided a Compaq laptop at the start of the field period.

Where CAPI Interviews Were Conducted

CAPI interviewers set appointments with NLS 12 respondents by the same means as those interviewers doing traditional paper and pencil interviews. CAPI interviews were conducted in the respondent's home, at the respondent's place of business, via telephone and, in several cases, in a penal institution. With one exception, CAPI interviewers were able to go wherever necessary to complete an interview using CAPI. In some cases CAPI interviewers did not have access to an alternating current power source and so made use of the rechargeable battery they were issued with each computer.

There was only one occurrence of an interview being refused because of CAPI: prison officials in one state refused to grant an NORC CAPI interviewer access to an incarcerated respondent because the contents of the machine were suspect, and could not be scanned by prison security personnel in advance.

How CAPI Interviews Were Completed

Each CAPI interviewer's laptop was preloaded in the central office with electronic face sheet data specific to the interviewer's caseload; the information on a given respondent's face sheet largely determined the default path followed during the interview. At the survey's conclusion, a special data backup and archiving program executed automatically, thus saving two backup copies of the completed interview (one version copied to the Compaq hard disk, another copied to a 3.5" floppy diskette) and preparing the case data for electronic transmission to the host computer in Chicago. Case data for interviews "broken off" before completion were not archived in preparation for electronic transmission.

How CAPI Case Data Were Sent to Chicago

CAPI interviewers sent completed case data to Chicago using special communication software loaded into their laptop computers. They were also required to send in weekly the 3.5" backup diskettes containing case data. The contents of the backup diskettes were compared, byte-by-byte, with case data transmitted electronically. Close analysis of data returned for every completed case revealed no differences between the two versions. Details of the computer support and transmission procedures are presented in Speizer's paper (1991).

Training

Seventy-seven men and women were trained as CAPI interviewers. Since these NORC interviewers were randomly selected to participate in the experiment, they possessed varying levels of skill and experience with both interviewing

and computer use. The training program was designed with this variety in mind. The aim was for complete flexibility: that is, to teach individuals at all experience levels. The training had three parts: self-study, in-person training, and post-training evaluation

Early in the field staffing process six of the NLS-12 Field Managers were selected to serve as CAPI field support during the field period. CAPI support training for field managers lasted one day and focussed on support of interviewers with CAPI difficulties. At the conclusion of this training, each of the CAPI support staff was issued a Compaq laptop computer, loaded with the CAPI software and data for 10 mock interviews. This enabled them to familiarize themselves with both the hardware and software. During this time Central Office CAPI training staff were available to answer any questions that arose among the support staff. Support staff familiarity with CAPI had the added advantage of allowing us to use them as assistant trainers during the CAPI interviewer training which took place two weeks after the support training. Members of the central office CAPI training staff were available to answer any questions that arose during this period.

Each interviewer was mailed a CAPI Self-Study Training Guide and instructed to spend 30 minutes completing it before arriving in Chicago for the training. Rather than focusing on the NLS interview itself, the guide focused on basic computer hardware and terminology. The goal was to ensure that all trainees arrived at the training with the same basic understanding of computer terminology. The guide proved extremely valuable in two ways. First, all trainees arrived at training with at least a basic level of computer knowledge although a few with previous computer experience possessed more than the others). Second, the descriptions and completed exercises served throughout the field period as valuable job aids. Furthermore, because of their familiarity with basic computer terminology, even the most inexperienced interviewers were able to accurately report any problems they encountered in the field.

Because interviewers were randomly selected to participate in the experiment, they came to the NLS with varying levels of interviewing experience, both in general and specifically on NLS. All CAPI interviewers received 3 days of NLS-CAPI specific training after completion of whatever general and study specific training was necessary because they lacked previous interviewing or NLS experience.

CAPI is a new technology, thus, we had available only a limited number of experienced CAPI trainers. This forced us to hold the CAPI training in two separate three day sessions. The first session contained all of the non-NLS experienced interviewers, as well as some experienced interviewers. The second session contained only NLS experienced interviewers.

We paid special attention to the needs of trainees for a high level of personal, hands-on experience with the CAPI program while they were still at training. The training was designed so that the lecture and demonstration modules were attended by the entire group. For the practice modules, the

group was divided into three smaller working groups and given hands-on exercises. This process enabled us to standardize the lecture and demonstration modules of the training while still giving each trainee personal attention.

Upon arrival at the training, each of the trainees received a set of CAPI materials designed to make their initial use of CAPI easier and less overwhelming. The most important of these materials was the CAPI Technical Manual which contained step-by-step instructions for performing any of the tasks asked of a CAPI interviewer. Thus, if interviewers encountered a problem with the CAPI software at any point during an interview, they could refer to this CAPI Technical Manual for the answer.

In addition to the CAPI Technical Manual, each of the interviewers received a set of CAPI administrative specifications, which outlined the procedures to follow when conducting a CAPI interview and a series of job aids to make the interviewing task easier.

The focus of training day 1 was helping interviewers overcome any initial computer fears they may have had. A brief session on how to pack and unpack the computer was held. (The main cause of damaged machines is the failure to pack them correctly for shipping.) This segment of the training had an added benefit when it became apparent that computer-novices found this familiarization time very reassuring. They were able to identify the various parts of the computer for the first time and gain ease in handling the equipment.

The rest of the first day was spent in a demonstration of the software, allowing the trainees their first hands-on experience with CAPI. The first hands-on session was conducted using a generic non-NLS questionnaire so that trainees could concentrate on the CAPI procedures without worrying about the specifics of the NLS instrument. The second half of the hands-on training consisted of practice with a mock NLS interview.

Day 2 covered the CAPI Administrative Specifications, which differed considerably from the paper-and-pencil procedures, an in-depth demonstration of the CAPI special features such as slide mode, fast forward, review of Face Sheet Information, etc., and more hands-on practice for the interviewers, highlighting the use of the special features.

The vast majority of Day 3 was spent doing more hands-on NLS practice. The trainees were required to complete an entire mock interview from the Household Enumeration to the Interviewer Remarks. The mocks were done in pairs, freeing up the trainers to move about the room and help the people who needed it most.

While trainees were completing these mock interviews, members of the training staff took pairs of interviewers and walked them through the transmission process. Interviewers were shown how to enter the transmission portion of the software, enter the appropriate codes and password, and transmit their completed cases. Also, at this time, interviewers had their actual NLS

caseload loaded into their computers, so that when they arrived home they were able to complete their telephone mocks and begin interviewing.

Day 3 concluded with interviewers repacking their computers and preparing them for shipment home.

Upon returning home from training, each interviewer was required to complete a telephone mock using the CAPI system. The telephone mocks were conducted by non-CAPI NLS Field Managers. This strategy allowed us to ensure that the FM was able to assist the interviewers to resolve any CAPI specific problems. This enabled us both to establish a test atmosphere and to model accurately an actual interviewing environment. Each of the seventy-seven interviewers was able to complete successfully the telephone mock and begin actual field interviewing.

We believe the CAPI training was a complete success. During the field period, only one CAPI interviewer, out of a total of 77, quit and that for reasons having nothing to do with CAPI. Only one required training beyond the initial training session. In comparison, the paper and pencil interviewers in the Control group lost 3 interviewers. It should be noted that all of the interviewers, who were randomly selected to participate in the experiment, attended the training. No one refused to work on NLS because of CAPI.

Operational Problems

In planning for the Round 12 NLS experiment NORC established procedures to address many problems that we expected to occur during the field period. While many of these problems did occur, others did not. Indeed, some problems we anticipated to be major turned out to be insignificant.

One of the problems we anticipated was that of developing a cost efficient way of providing support to interviewers during "non-business" hours. The first step in developing this support structure was to establish a field support network. Since the computer programmers' time is very expensive, we wanted to ensure they were involved with problems only they could resolve. The field support network was designed to screen calls and resolve as many problems as possible without programmer intervention.

In anticipation of problems solvable only by a programmer, however, an NORC programmer was "on-call" during non-business hours for the first few weeks of the field period. We found the number of calls requiring programmer intervention rapidly dwindled, and, by the 4th week, we ceased to need programmer assistance. At this point, the field support network took on more responsibility in diagnosing problems and soon became quite proficient. If the field support network could not resolve the problem, a member of the NLS staff was contacted at home. This occurred only three times during the entire field period.

NORC developed a plan for quick replacement of field computers. Replacement computers were always sent within 24 hours. If the problem occurred over a

weekend, the replacement procedures were implemented first thing Monday morning. The number of hardware problems experienced during the field period, however, was far fewer than originally expected. In fact, only five machines broke down and needed to be replaced; three of these problems were diagnosed during the training.

Interviewer Errors

Due to the quick receipt of the CAPI cases, NORC was able to detect, analyze and resolve interviewer errors much more effectively than those detected in PAPI questionnaires. When interviewers encountered a problem in CAPI, they were instructed to describe what happened in a comment field available at any point during the CAPI interview. Additionally, in the Interviewer Remarks section of the CAPI questionnaire, interviewers were asked if they had encountered any problems. If they responded positively, they were asked again to describe the problem in detail. Whenever a "Yes" was recorded in response to this question, CAPI automatically flagged the case as a "problem case" requiring instant editing.

Problem cases were sent to members of the CAPI staff, usually the day after they were received. Staff reviewed the case to determine what happened and how to resolve it. In many cases, the errors reported were very simple things, e.g., the interviewer forgot the password to get into the case and had to look it up in the manual. In other cases, they pointed to more serious errors. The vast majority of problems encountered involved difficulty in entering/editing rosters. A combination of changes to the software and improved training procedures should minimize this type of error in the future.

Transmission Problems

Before beginning the discussion of transmission problems, it is important to note that transmission was not considered to be a success unless all parts of the transmission were completed successfully. Thus, if an interviewer attempted to transmit three cases, and was successful with the transmission of only two, the transmission was considered to be a failure, even though two cases were successfully transmitted and received.

Success rates in the early part of the field period (July 1990) averaged 80%. At that time NORC instituted some changes in the transmission software that increased the success to 90% on the first transmission try, where it remained for the rest of the field period. The vast majority of the transmission problems that remained were due to bad telephone line connections. If the interviewer terminated the connection and attempted to transmit again, the problem was solved. It should be noted that the communication software prohibited the dispatch of incomplete case data.

SYSTEMS TO SUPPORT CAPI

The NORC Field Management System (FMS)

In an effort to facilitate the management of the CAPI experiment, NORC both refined many of its existing project monitoring systems and developed new ones. The most notable of these systems was the Field Management System, (FMS), which was used for the first time on NLS during Round 12. In addition to allowing Field Managers the ability to monitor the progress of their individual regions electronically, on a case-by-case basis, the system allowed the central office and the field managers to share data in new ways. For instance, the field managers' computers were updated weekly as cases were received in the central office. When a CAPI case was electronically received in the office a code was returned to the field manager, through the FMS, to indicate that the case had been transmitted and successfully received.

One of the biggest advantage of using the FMS was its ability to enforce the basic sample control rules outlined earlier. The FMS was designed so that illegal assignments, e.g. assigning Control cases to a CAPI interviewer, were not allowed. Thus, by using the FMS, we were able to ensure the integrity of the experiment.

Making "Fixes" to the Software

In some cases, early in the field period, errors in the questionnaire software were identified and the need for a fix was established. These errors paralleled the type that are traditionally corrected with errata memos to the interviewers. CAPI, however, offered an opportunity to correct these errors centrally, thus ensuring that each interviewer was following the correct procedures.

During Round 12 five fixes to the CAPI questionnaire were needed, all within the first few weeks of the field period. When it was determined that a fix was necessary, interviewers were notified and instructed to dial in to receive the fix electronically.

Respondent Opinion of CAPI

Each CAPI respondent was asked to complete a CAPI Respondent Feedback Questionnaire at the conclusion of the Main interview. To date, 1,131 questionnaires have been partially analyzed. The preliminary analysis indicates that respondent opinion about CAPI on NLS is generally favorable, often enthusiastically supportive. Most of the CAPI respondents stressed that this year's survey, administered via the laptop computer, seemed easier for the interviewer, was less stressful than in the past, i.e., than with paper & pencil, and seemed more accurate.

Cost of CAPI Cases

The cost elements for CAPI and PAPI cases are differently distributed, with training costs being higher and data processing costs lower for CAPI as compared with PAPI. These are cost differences that will persist by mode even

after the method is firmly established and frequently used, although CAPI training costs may come down somewhat as more interviewers become generally familiar with computers. We estimate that the 3 days devoted to CAPI training might be reduced to 2.5 if all interviewers were experienced computer users and did not have to be familiarized with computers in general before beginning training on the particular CAPI system and the particular study.

At prevailing prices for laptops at the time we conducted this experiment, the total costs of the CAPI and PAPI cases were almost exactly the same. The training costs ran about 69% higher for CAPI than for PAPI, the field costs for interviewing labor were about 1% higher for CAPI, but if you add the cost of the laptops (amortized over 3 years) to the interviewing costs, then interview costs for CAPI were about 38% higher than PAPI. On the other hand, data processing costs were almost 500% higher for PAPI than for CAPI cases which almost exactly offset the greater CAPI costs of the other elements. Thus, the total direct costs for the two modes in this experiment were almost exactly the same. These costs do not reflect the costs of developing and fixing the software and other CHRR costs nor the costs of administering the experiment itself.

Direct interviewing costs for CAPI-administered cases were appreciably more expensive during the first two months of the field period, costing an average of $8.00 more per completed case than the PAPI control case through Week 12 of the field period. After Week 12, the cost of completing cases via CAPI decreased while the cost to complete cases using traditional materials rose slightly; by Week 17, the last week of the CAPI experiment, the cumulative cost per completed case was only 1% higher than that of the Control sample.

We must be cautious in generalizing this cost experience to other surveys since the final cost difference will depend on the mix of the elements in a particular survey. For NLS, data processing costs are very large relative to other costs, thus considerable savings in this element was achieved by using CAPI, and those savings offset the extra expenses for training and the cost of machines.

A big cost factor, of course, is the cost of the laptops, but the cost of acceptable hardware is declining rapidly which will help make CAPI more cost effective in the future. The depreciation period used and how one allocates machine costs across studies will also affect the costs for a particular study. For this experiment, we assumed a machine life of 3 years and, because of the lengthy field period, we assigned all of the machine costs to this study. Other studies with a different mix of interview complexity, assignment size and data processing complexity might have a different cost comparison between the two modes.

CONCLUSION

We conclude that CAPI is a feasible method to use for large scale national surveys even with present technology. For large, complex surveys it is

probably cost effective at present hardware prices if the machine costs can be spread over a number of years and the machines can be used at least 50% of the time. The logistical problems in maintaining the machines, training and supporting the interviewers and in transmitting the data electronically are all well within our capabilities. The data concerning the effect of CAPI on data quality or possible mode differences in responses are still to be analyzed, but so far we are optimistic that there will be few problems, and many advantages.

REFERENCES

DENTENEER, D., BETHLEHEM, J.G., HUNDERPOOL, A.J., AND KELLER, W.J. Computer Assisted Survey Processing," in Proceedings of the Third Annual Research Conference, Washington, DC: U.S. Bureau of the Census, 112-127.

ROTHSCHILD, BETH B. AND WILSON, LUCY B.(1988), "Nationwide Food Consumption Survey, 1987: A Landmark Personal Interview Survey Using Laptop Computers," in Proceedings of the Fourth Annual Research Conference, Washington,DC: U.S. Bureau of the Census, 347-356.

SEBESTIK, JUTTA, ZELON, HARVEY, DEWITT, DALE, O'REILLY, JAMES AND McGOWAN, KEVIN (1988), "Initial Experiences with CAPI," in Proceedings of the Fourth Annual Research Conference, Washington,DC: U.S. Bureau of the Census, 357-365.

SPEIZER, HOWARD AND DOUGHERTY, DOUG (1991), "Automating Data Transmission and Case Management Functions for a Nationwide CAPI Study," presented at the U.S. Bureau of the Census Annual Research Conference, Arlington, VA, 1991.