

Evaluating the Computer Audio-Recorded Interviewing (CARI) Household Wellness Study (HWS) Field Test

Taniecea A. Arceneaux
U.S. Census Bureau

Abstract

This article provides the results of research, testing, and analysis conducted by the U.S. Census Bureau. As a supplement to traditional reinterview, Computer Audio-Recorded Interviewing (CARI) will allow more efficient sampling methods and a more focused reinterview program while maintaining the Census Bureau's standard of data quality. With the respondent's consent, CARI unobtrusively recorded the verbal exchange between the interviewer and the respondent without disrupting the normal interview process.

While the broader goal of this research is to implement CARI into all of the Census Bureau's computer-assisted personal interview (CAPI) surveys, the present study accomplishes the following tasks: evaluate the impact of CARI on data quality, measure the audio quality of recordings collected during personal visit interviews, determine the impact of CARI on system performance, and obtain the reactions of respondents and Field Representatives (FRs) to the use of CARI.

Keywords: Computer Audio-Recorded Interviewing, Quality Assurance, Computer Assisted Personal Interviewing

1. Background

Computer Audio-Recorded Interviewing (CARI) is a laptop software application, developed by the Research Triangle Institute (RTI). With the respondent's consent, CARI unobtrusively records the verbal exchange between the interviewer and the respondent without disrupting the normal interview process. By reviewing the recorded portions of the interview, quality assurance monitors can evaluate the likelihood that the exchange between the FR and the respondent is authentic.

To demonstrate the feasibility of using CARI, previous research supported by the RTI and the U.S. Department of Health and Human Services has shown the following:

- CARI processes are feasible for production surveys.

- Adequate audio quality can be achieved to allow both the FR and respondent to be heard clearly and distinctly.
- Reviewing three 30-second audio files is sufficient for monitors to reach a consensus on the validity rating of the interview.
- CARI files are compatible with the Entrust encryption system.
- CARI is compatible with Blaise.
- Using file compression methodology, it is feasible to transmit CARI files via 56K modems at an acceptable rate.

Despite these positive findings, more research is necessary before CARI can be implemented for use in Census Bureau current surveys. The recent experience with CARI made it desirable to conduct a series of lab and field tests. In 2004, the CARI team conducted two lab tests to test CARI technology applied to a CAPI Blaise instrument. These tests were conducted using Census Bureau Headquarters (HQ) staff as interviewers and monitors. The audio files were manually removed from the laptops and manually compressed and encrypted. During these lab tests, the following have been shown:

- The interviewer cannot detect the recorder.
- CARI is compatible with Blaise.
- The recorder can be controlled by inserting code in the instrument.
- CARI code does not increase the length of the interview.
- Audio files are of adequate sound quality (no adverse affect from compression and encryption).

The CARI HWS Field Test was designed to further test the feasibility of CARI in the field. This research program is being implemented jointly by the Demographic Statistical Methods Division (DSMD), the Demographic Surveys Division (DSD), the Field Division (FLD), and the Technologies Management Office (TMO).

2. Data Source

2.1 Instrument Design

The National Health Interview Survey (NHIS) instrument was chosen to be used as a base for the

CARI HWS Field Test 1. The 2005 Quarter 2 NHIS instrument was modified to create a Household Wellness Study instrument sponsored by the Census Bureau. The instrument changes included turning off the Spanish translation, removing the cancer supplement, changing all references from NHIS to “Health Wellness Study,” adding verbal and recorded consent questions to obtain recording consent from each respondent in the survey, adding respondent debriefing questions, and inserting code to control the functionality of the recorder. With consent of each respondent, there were up to three 30-second portions of the interview recorded.

2.2 Sample Selection

The DSMD selected the CARI HWS sample from the 23 cancelled weeks of the 1990 NHIS redesign sample. The sample was selected from the 1990 Sample Control Database. The cases were selected from 71 Primary Sampling Units (PSUs) that are also in the 2000 sample redesign so that current NHIS staff would be available to complete the interviews. The sample of 828 HWS cases was selected from the Philadelphia, Detroit, and Kansas City Regional Offices (ROs).. During the CARI HWS Field Test 1, FRs from each of those three ROs conducted the interviews.

Since the cases for the CARI HWS Field Test 1 were selected from PSUs that were in both the 1990 and 2000 sample redesign, there was a possibility that a case selected in the 1990 NHIS redesign sample was also selected in the 2000 NHIS redesign sample. A household could potentially be interviewed twice, once for the CARI HWS Field Test 1 and again for the 2006-2010 NHIS. Therefore, an unduplication process was completed before the sample selection took place.

3. Research Questions

3.1 Impact of CARI on Data Quality

3.1.1 Do the CARI files corrupt production data?

The DSMD determined from the Technologies Management Office (TMO) Technical Assistance Center (TAC) if there was corruption of production data. If the production data from less than 4 percent of the cases were corrupted, then the CARI corruption rate is deemed acceptable. The formula used for calculating the CARI corruption rate is as follows:

$$\text{CARI Corruption Rate} = \frac{\text{Number of cases with corruption of production data}}{\text{Total number of cases eligible for interviewing}} \times 100$$

According to the TMO, there was no indication of the corruption of production data. Therefore, the CARI corruption rate is 0 percent. The data indicates that CARI files do not cause corruption of production data.

3.1.2 What are the CARI HWS Field Test response rates?

The DSMD provided and analyzed measures of response rates and refusal rates by RO and in total for both the CARI HWS Field Test and the corresponding weeks of the 2006 NHIS. The CARI HWS Field Test response rate is defined as the ratio of the number of responding units (complete and sufficient partial interviews) to the total number of cases eligible for interviewing. The CARI HWS Field Test refusal rate is defined as the ratio of the number of refusals to the total number of cases eligible for interviewing. The total number of cases eligible for interviewing used in both rates is calculated as the total number of cases selected minus the Type B and C cases (ineligibles). The response rates and refusal rates were calculated without weighting the data. Table 3.1.2.1 below outlines the response rates and refusal rates by RO and in total for each of the three regions participating in the CARI HWS Field Test.

Table 3.1.2.1 CARI HWS Field Test Response Rates and Refusal Rates by Regional Office

	Response Rate	Refusal Rate
Total	81.41	15.34
Philadelphia	76.70	18.75
Detroit	83.33	12.78
Kansas City	83.84	14.65

Since the CARI HWS instrument was a modified NHIS instrument, the response rates for the 2006 NHIS cases (from Quarter 1, Weeks 5-13 and Quarter 2, Weeks 1-2) completed in the same time frame, namely from January 30, 2006, to April 16, 2006, were compared to the CARI HWS response rates for the interviews completed from February 1, 2006, to April 11, 2006. A major difference in the way the CARI HWS was conducted is that there were no screening rules applied to the sampling procedure for households in the field test. That is, the 2006 NHIS implemented a procedure for screening certain households out of the survey depending on their race and ethnicity. Therefore, to make the response rates comparable, the NHIS response and refusal rates were recalculated including cases that would have screened out.

The NHIS response rate is defined as the ratio of the number of complete and sufficient partial interviews and screened-out cases in the 2006 NHIS to the total number of cases eligible for interviewing. The NHIS refusal rate is defined as the ratio of refusals to the

total number of cases eligible for interviewing. The total number of cases eligible for interviewing used in both rates is calculated as the total number of cases selected minus the Type B cases (excluding screened-out cases) and the Type C cases. The response rates and refusal rates were calculated without weighting the data. Table 3.1.2.2 below outlines the response rates and refusal rates by RO and in total for the corresponding weeks of the 2006 NHIS for each of the three regions participating in the CARI HWS Field Test. Due to restrictions on allowed sample size, CARI interviewing was immediately terminated when at least 423 CARI cases were completed.

Table 3.1.2.2 2006 NHIS Response Rates and Refusal Rates by Regional Office

	Response Rate	Refusal Rate
Total	90.09	6.54
Philadelphia	84.84	9.30
Detroit	92.95	5.70
Kansas City	91.78	4.98

Hypothesis testing was used to compare the overall response rates and refusal rates for the CARI HWS Field Test and the 2006 NHIS. One-tailed z-tests were used to compare proportions at the 10-percent level of significance. In testing whether the CARI HWS Field Test response rate is significantly less than that of the 2006 NHIS, the p-value is very close to 0 at the 10-percent significance level. That is, the overall response rate for the CARI HWS Field Test is significantly less than the overall response rate for the 2006 NHIS. In testing whether the CARI HWS Field Test refusal rate is significantly higher than that of the 2006 NHIS, the p-value is very close to 0 at the 10-percent significance level. That is, the overall refusal rate for the CARI HWS Field Test is significantly higher than the overall refusal rate for the 2006 NHIS.

3.2 Audio Quality of CARI Recordings

3.2.1 What is the quality of the audio recordings collected during personal visit interviews?

Each time an FR transmitted cases, the .wav files were compressed, zipped, encrypted, and transmitted with the production data. At Headquarters, the files were decrypted and unzipped to extract the .mp3 files. Headquarters staff acted as monitors and listened to the .mp3 files using headphones. The monitors listened to all of the recorded consent files (hereafter referred to as CARICON files) and the 30-second recorded portions of the interview (hereafter referred to as snippet files). They evaluated the recording quality of each CARICON and snippet file using the five-point scale rating system provided in Table 3.2.1.1 below.

Table 3.2.1.1 Audio Quality Rating Scale

Rating	Description
Excellent	Both the interviewer and respondent are clearly intelligible
Good	One can clearly hear the interviewer and determine that a respondent is participating, but some of the respondent's reply is unintelligible.
Fair	One can clearly hear the interviewer, but the respondent cannot be heard at all.
Poor	Neither the interviewer nor the respondent is intelligible (e.g., the recording consists of white noise).
Too Short	The audio file is less than 5 seconds in length.
Cannot Rate	One cannot evaluate the quality of the audio file because of some technical problem (e.g., the file cannot be opened).

Approximately once a week (or daily by the end of the interviewing period), the DSMD received the zipped audio files and audit trail files for the cases received since the last transmission. The monitors listened to each audio file and completed an assessment form for each interview, detailing the audio quality rating for each audio file obtained. The results of this assessment can be found in Table 3.2.1.2 below, where the first percentage shown in each cell is the row percent, and the second percentage shown in each cell is the column percent.

Table 3.2.1.2 Audio Quality Ratings of CARI Audio Files by Type—Frequency and Percent

	CARICON	Snippet	Total
Excellent	390	1078	1468
	(63.8%)	(89.5%)	(80.9%)
Good	(26.6%)	(73.4%)	(100.0%)
	18	67	85
Fair	(2.9%)	(5.6%)	(4.7%)
	(21.2%)	(78.8%)	(100.0%)
Poor	95	55	150
	(15.5%)	(4.6%)	(8.3%)
Too Short	(63.3%)	(36.7%)	(100.0%)
	12	4	16
Cannot Rate	(2.0%)	(0.3%)	(0.9%)
	(75.0%)	(25.0%)	(100.0%)
Total	96	0	96
	(15.7%)	(0.0%)	(5.3%)
Total	(100.0%)	(0.0%)	(100.0%)
	0	0	0
Total	(0.0%)	(0.0%)	(0.0%)
	(0.0%)	(0.0%)	(0.0%)
Total	611	1204	1815
	(100.0%)	(100.0%)	(100.0%)
Total	(33.7%)	(66.3%)	(100.0%)

The DSMD summarized the results of the monitors' evaluations. If more than 96 percent of the recordings were rated Excellent or Good, then the audio quality is deemed acceptable. The formula for calculating this percentage is as follows:

$$\text{High-Quality Audio File Rate} = \frac{\text{Number of audio files rated "Excellent" or "Good"}}{\text{Total number of audio files (CARICON and snippets)}} \times 100$$

Since the observed high-quality audio file rate is 85.6 percent, the audio quality is deemed unacceptable. In examining the source of this unacceptability, a few noteworthy findings are important. Training issues may account for the large percentage of fair CARICON files. Further analysis of the CARICON files revealed that in 77.9 percent of CARICON files rated "Fair" (74 out of 95), the respondent's reply could not be heard because the FR keyed the CARICON response before the respondent's consent was recorded by the instrument. FRs must be trained to allow the respondent to verbally reply to the CARICON question before keying the response into the instrument, as this will turn off the recorder.

Examining the definition of the audio quality rating system also indicates a possible cause of the unacceptable audio quality. The audio quality rating system was designed to measure the audibility of both the interviewer and the respondent's voice on the recording. However, when taking interview situations into consideration, the rating system may not have captured the true audio quality of the audio files. For instance, in the case in which a snippet recording contained only keying sounds, the file was rated Poor (no voices could be heard at all), but the actual audio quality of the recording may have been adequate. Therefore, it is desirable to redesign the audio quality rating system in order to gain more precise results of audio quality before a definitive acceptability decision can be made.

3.3 Impact of CARI on System Performance

At the end of each interview, the FRs were asked a series of hardware performance questions in the instrument. The purpose of these questions was to determine if CARI had an effect on instrument operation throughout the course of the interview. The FRs were also asked a series of hardware performance questions at the conclusion of the interviewing period in an FR debriefing questionnaire.

$$\text{System Malfunction Rate} = \frac{\text{Number of cases indicating technical problems}}{\text{Total number of cases eligible for interviewing}} \times 100$$

3.3.1 What is the impact of CARI on hardware performance?

The DSMD tallied the responses to the hardware performance questions in the instrument. If the interviewers indicated that he/she noticed technical problems in less than 4 percent of the cases eligible for interviewing, then the system malfunction rate is deemed acceptable. The formula for calculating the system malfunction rate is as follows:

In response to the questions in the instrument, the FRs indicated that they experienced technical problems in 8 of the 447 cases. Therefore, the system malfunction rate is 1.8 percent. The data indicates that CARI does not have an impact on hardware performance.

3.3.2 Are the interviewers able to detect the recorder turning on and off?

The DSMD tallied the responses to the hardware performance questions in the instrument and the FR debriefing questionnaire. If the interviewers indicated that he/she could detect the recorder turning on or off in less than 4 percent of the cases eligible for interviewing, then the recorder detection rate is deemed acceptable. The formula for calculating the recorder detection rate is as follows:

$$\text{Recorder Detection Rate} = \frac{\text{Number of cases indicating recorder detection}}{\text{Total number of cases eligible for interviewing}} \times 100$$

In response to the questions in the instrument, the FRs indicated that they could detect the recorder turning on or off in 3 of the 447 cases. Of the 3 cases in which the FR did indicate detection of the recorder, all 3 were completed by the same FR. Upon further inspection of that FR's debriefing questionnaire, it is evident that the interviewer was simply confused by a flashcard indicator on the instrument screen. For this reason, the 3 cases that indicated detection of the recorder were the result of training issues and are not included in the recorder detection rate. Therefore, the recorder detection rate is 0 percent. The data indicates that the interviewers are not able to detect the recorder turning on and off.

3.3.3 Does the recorder function properly?

The DSMD determined if the CARI recorder functions properly by reviewing the original interview data and comparing the number of recordings expected to the number of recordings received. The criteria for determining if the recorder functions properly are the following:

- There is a CARICON recording each time there is a “Yes” response to the verbal consent question during the interview.
- There is no CARICON recording if there is a “No” response to the verbal consent question during the interview.
- If the respondent says “Yes” to the recorded consent question at the beginning of the interview and does not change his response to “No” at any point during the interview, then there should be three 30-second snippets recorded. In the case of a partial interview, there may be less than three snippets recorded.
- If the respondent says “No” to the recorded consent question at the beginning of the interview and does not change his response to “Yes” at any point during the interview, then there should be no snippets recorded.
- If the respondent says “Yes” to the recorded consent question at the beginning of the interview and changes his response to “No” at any point during the interview, then there may be up to three 30-second snippets recorded. (Depending on skip patterns in the instrument, there may be less than three snippets recorded.)
- If the respondent says “No” to the recorded consent question at the beginning of the interview and changes his response to “Yes” at any point during the interview, then there may be up to three 30-second snippets recorded. (Depending on skip patterns in the instrument, there may be less than three snippets recorded.)

If less than 4 percent of the cases failed at least one of the above criteria, then the recorder malfunction rate is deemed acceptable. The formula for calculating the recorder malfunction rate is as follows:

$$\text{Recorder Malfunction Rate} = \frac{\text{Number of cases with a recorder malfunction}}{\text{Total number of cases eligible for interviewing}} \times 100$$

The number of recordings expected from the original interview data was equal to the number of recordings received. That is, there were no cases in which the recorder malfunctioned. Therefore, the recorder

malfunction rate is 0 percent. The data indicates that the recorder functions properly.

3.4 Respondent and FR Reactions to CARI

3.4.1 What are the CARI cooperation rates for the CARI HWS Field Test?

The DSMD calculated the CARI cooperation rates from the responses to the verbal and recorded consent questions in the CARI HWS Field Test interviews. The CARI cooperation rates are defined as follows:

$$\text{Full CARI Cooperation Rate} = \frac{\text{Number of households who agreed to be recorded for the entire interview}}{\text{Number of households who were asked the CARICON question}} \times 100$$

$$\text{Partial CARI Cooperation Rate} = \frac{\text{Number of households who agreed to be recorded for part of the interview and refused for part of the interview}}{\text{Number of households who were asked the CARICON question}} \times 100$$

$$\text{Full CARI Refusal Rate} = \frac{\text{Number of households who refused to be recorded for the entire interview}}{\text{Number of households who were asked the CARICON question}} \times 100$$

Table 3.4.1.1 below outlines the CARI cooperation rates by RO and in total. It lists the number of cases by RO in which the household agreed to be recorded for the entire interview, agreed to be recorded for part of the interview and refused to be recorded for other parts of the interview, or refused to be recorded for the entire interview. The percentages shown in each cell represent the row percent. Standard errors of these sample proportions are also calculated.

Table 3.4.1.1 CARI Cooperation Rates by Regional Office – Frequency and Percent

	Full CARI Coop	Partial CARI Coop	Full CARI Ref	Total
Total	414 (88.7%)	14 (3.0%)	39 (8.4%)	467 (100%)
Philadelphia	119 (83.8%)	4 (2.8%)	19 (13.4%)	142 (100%)
Detroit	139 (89.1%)	7 (4.5%)	10 (6.4%)	156 (100%)
Kansas City	156 (92.3%)	3 (1.8%)	10 (5.9%)	169 (100%)
Std Errors	0.015	0.008	0.013	

3.4.2 What are the respondents' reactions to CARI?

The FRs asked the respondents a series of debriefing questions at the end of each interview regarding the use of CARI. The DSMD reviewed and tallied the responses to the respondent debriefing questions.

When asked to elaborate about their most recent consent response in the interview, 3.3 percent of respondents (14 out of 430) provided an elaboration detailing a negative reaction to being asked to be recorded. Of those 14 respondents, 8 cited privacy concerns, 5 were simply uncomfortable with being recorded, and 1 did not understand the purpose of the recordings.

When asked if they had any concerns about what would happen with their answers if they were recorded, 88.5 percent of the respondents who provided consent (332 out of 375) did not have any concerns. Of the 43 respondents who did have concerns with recording, 36 of them expressed specific issues. Of those 36 respondents, 20 had privacy concerns, 14 were unsure about the use of the recordings, and 2 expressed anti-government sentiments.

At the conclusion of interviewing, the FRs completed debriefing questionnaires that contained questions relating to respondent reactions. The DSMD reviewed and tallied the responses to the FR debriefing questionnaires.

Of the 44 interviewers who completed the debriefing questionnaire, only 16 of them reported encountering respondents who were reluctant to be recorded. The FRs who encountered reluctance reported using the following techniques to successfully persuade respondents to agree to CARI: developing rapport (i.e., light conversation, politeness), emphasizing confidentiality, suggesting that the respondent may refuse to answer a particular question that may be too personal, and reiterating the purpose of recording (i.e., quality control, reporting accurate information, interview verification).

3.4.3 What are the NHIS FRs' experiences and concerns with CARI?

The FR debriefing questionnaire also contained questions relating to FR reactions to CARI. A series of statements was provided, and the interviewers completing the form were asked to rate their agreement with each statement. The DSMD reviewed and tallied the responses to the FR debriefing questionnaires.

FRs had mixed reactions to the introduction of CARI into the production environment. When they first learned of CARI, 59.1 percent of interviewers (26 out of 44) felt neutral or comfortable with the fact that their interviews were being audio recorded, but by the time interviewing was over, this figure had increased to 65.9 percent (29 out of 44). One of the FRs originally comfortable or neutral with CARI changed to being uncomfortable with CARI when the interviewing was over. But 6 of the 19 FRs originally uncomfortable with CARI changed to being either neutral or comfortable with CARI when interviewing was over.

When completing the debriefing questionnaires, 61.4 percent of interviewers (27 out of 44) understood the reasons why their interviews were being recorded, and 52.3 percent (23 out of 44) agreed that recording interviews is a better way to verify the authenticity of interviews and evaluate interviewer performance. However, only 47.7 percent of them (21 out of 44) agreed that they personally felt very positive about the use of audio recording as a way to verify that their interviews have actually been conducted.

In addition, 54.5 percent of interviewers (24 out of 44) agreed that having monitoring staff listen to audio recordings of interviews is a good way to evaluate an employee's performance as an FR, but only 36.4 percent of them (16 out of 44) felt positively about the use of audio recording of interviews as a way to evaluate their personal performance and provide feedback.

4. Limitations

To give proper interpretation to the results presented in this article, it is important to keep in mind the following limitations.

4.1 Limited Universe

The universe of this field test was restricted to only three regional offices, namely Philadelphia, Detroit, and Kansas City. Therefore, there is no data available to evaluate or predict the response rates, CARI cooperation rates, respondent reactions, and FR reactions for CARI surveys in other regions.

4.2 No Screening Rules

The CARI HWS instrument was a modification of the 2005 Quarter 2 NHIS instrument. Comparisons were made in this report between the response rates and refusal rates for the CARI HWS and the 2006 NHIS, which are very similar surveys completed by the same regional offices in the same time period. A major

difference in the way the CARI HWS was conducted is that there were no screening rules applied to the sampling procedure for households in the survey. That is, the 2006 NHIS implemented a procedure for screening certain households out of the survey depending on their race and ethnicity. Therefore, to make the response rates comparable, the NHIS response and refusal rates were recalculated including cases that would have been screened out.

4.3 Training Issues

Training issues may have contributed to the unacceptable audio quality ratings for the recordings captured in this test. With the knowledge gained through this analysis, future tests will allow better training to occur, which would possibly increase the audio quality of recordings.

4.4 Lesser Sense of Urgency

The nature of a field test invokes a lesser sense of urgency than normal production work. The FRs completing the interviews for the CARI HWS Field Test also worked on production work for the NHIS. These FRs were instructed that their regular production work was to take priority over completing cases for the field test. Therefore, the response rates in the CARI HWS Field Test are not an adequate indication of what to predict in a CARI production survey.

5. Conclusions

5.1 Impact of CARI on Data Quality

The CARI HWS Field Test has revealed many important implications for the incorporation of CARI into the Census Bureau's current quality assurance program. Regarding the impact of CARI on data quality, the field test data indicates that the CARI audio files do not cause the corruption of production data. As an indirect measure of data quality, the response rates of the CARI HWS Field Test were significantly lower than that of the 2006 NHIS, but the fact that the CARI HWS was in fact a field test may have impacted the urgency in the FRs' efforts to obtain high response rates.

5.2 Audio Quality of CARI Recordings

The data also indicates that 85.6 percent of all audio files recorded in the field test were rated either Excellent or Good. According to Section 3.2, an acceptable high-quality audio file rate is 96 percent. Therefore, the audio quality in this test is deemed unacceptable, and the most likely cause lies in training issues, especially in regards to allowing the respondent

to verbally respond to the CARICON question before keying their response and pressing the Enter key (which turns off the recorder). Furthermore, redesigning the audio quality rating scale would allow more precise analysis of true audio quality.

5.3 Impact of CARI on System Performance

With the introduction of any new technology comes the possibility of an effect on system performance. As a result of the CARI HWS Field Test, the data indicates that the recorder functions properly, and the recorder turning on and off is undetectable by both the respondents and the FRs. Furthermore, CARI does not appear to adversely affect the incidence of technical problems during CAPI surveys.

5.4 Respondent and FR Reactions to CARI

Respondents typically were very receptive to the introduction of CARI in this production environment. Of all households who were asked the CARICON question, 88.7 percent accepted CARI for the entire interview, and 3.0 percent accepted CARI for part of the interview. Those respondents who did not agree to be recorded mentioned privacy concerns, anti-government concerns, and uncertainty about the use of the recordings as their reasons for refusal.

Field Representatives have mixed reactions to the introduction of CARI into the production environment. Some FRs (38.6 percent) were comfortable with the fact that their interviews were recorded, but some (22.7 percent) were adamantly opposed to CARI. Generally, FRs understood the reasons why their interviews were being recorded and agreed that recording interviews is a better way to verify the authenticity of interviews and evaluate interviewer performance than the current non-CARI system. Some FRs (36.4 percent) believed that audio recording of interviews keeps them focused and improves their performance, while others (29.5 percent) believed that the recording was a deterrent.

6. Recommendations for Future Research

In order to fully integrate CARI into the production of all current Census Bureau surveys, several systems must be designed and tested in a future dress rehearsal. The next test will focus on the development of a monitoring system for the evaluation of interviewer performance and interview authenticity, as well as the development of a quality assurance sample design which incorporates CARI into the current reinterview program. This test should be designed using a survey other than the NHIS in order to measure respondent and FR reactions to the use of CARI in surveys collecting different types of information than was

collected in the CARI HWS. In addition, measures of data quality and system production will also be evaluated.

Because the response rates and refusal rates obtained in the CARI HWS Field Test were potentially impacted negatively by the fact that FRs were operating in a testing situation, it is important to analyze such measures of data quality using data obtained in a production survey setting, which will compare two treatments—one receiving the CARI treatment and the other without the CARI treatment.

The audio quality of the CARI audio files obtained and monitored through this field test is deemed unacceptable. The cause of this unacceptability is a combination of inadequate interviewer training and an inconclusive audio file rating system. A future CARI test would allow more efficient interviewer training and a more precise monitoring system, as a result of the knowledge gained through the CARI HWS Field Test.

The results of the CARI HWS Field Test indicate that the FRs have mixed reactions regarding the use of CARI for quality assurance purposes. In future CARI tests, it is highly desirable to determine a method of introducing the CARI technology to interviewers in such a way that would obtain more support and positive reactions from FRs.

Acknowledgements

The work presented in this article resulted from the contributions of many individuals.

In the planning phase of the CARI HWS Field Test, a team consisting of members from the Demographic Statistical Methods Division, the Demographic Surveys Division, the Technologies Management Office, and the Field Division was formed to design the methodology for implementing the field test. The CARI team consisted of the following members: Cynthia Wrenn-Yorker, Candice Barnes, David Montgomery, Patricia Hartman (project manager), Taniecea Arceneaux, and Phyllis Singer of the Demographic Statistical Methods Division; Heather Holbert of the Demographic Surveys Division; Byron Crenshaw and Mark Gorsak of the Field Division; and Malcolm Robert Wallace of the Technologies Management Office. Robyn Sirkis, now in the Decennial Statistical Studies Division, was also a member of the CARI team in its initial stages.

During the implementation of the CARI HWS Field Test, several other people joined the core team in order to produce the test in the field. These additional team

members included the following: Mircea Ardelean, Christine Freeland, Jeffery Garnto, Li Li, Sean McLaren, David Skeeahan, Lynn Swindler, Charles Tyler, Jr., Betty Watkins, and Sadie Whitfield of the Technologies Management Office; Patricia Holanek, Deloris Cockrell, Caroline Evans, and Jeff Wneck of the Demographic Statistical Methods Division; Jeff Behler and Roderick Wiley of the Field Division; Thomas Almerini, Fernando Armstrong, Philip Lutz, and Theodore Roman of the Philadelphia Regional Office, Tammy Battiste and Raef Fakhoury of the Detroit Regional Office, and Dennis Deeney of the Kansas City Regional Office.

Many other people throughout the Census Bureau also contributed to the advancement of this project.

References

- Arceneaux, T. "Evaluation Report for the Computer Audio-Recorded Interviewing (CARI) Household Wellness Study (HWS) Field Test 1," Bureau of the Census, August 15, 2006.
- Hartman, P., "Instrument Specification for the Computer Audio Recorded Interviewing Field Test #1 (2006 Household Wellness Study)," Bureau of the Census, 2005.
- Singer, P. and Sirkis, R., "Evaluation Plan for CARI Field Test 1," Bureau of the Census, 2005.
- Singer, P. and Sirkis, R., "Revised CARI Project: Specifications for Area Frame Sampling for CARI Field Test 1 Using the National Health Interview Survey (NHIS) Sample," Bureau of the Census, November 30, 2005.
- Singer, P. and Sirkis, R., "Specifications for the Unduplication of the Sample for the Computer Audio-Recorded Interviewing (CARI) Field Test 1," Bureau of the Census, December 20, 2005.

This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on statistical, methodological, technical, or operational issues are those of the author and not necessarily those of the U.S. Census Bureau.