# Computer Audio Recording: A Practical Technology for Improving Survey Quality

M. Rita Thissen[1], Hyunjoo Park[1], Mai Nguyen[1]

[1] RTI International

In interviewer-administered surveys, some aspects of total survey error (Biemer, 2010) can be managed through use of computer audio-recorded interviewing (CARI) technology by addressing faulty question wording or translation, mode effects, inconsistent question presentation, field data entry errors, and some aspects of item non-response. CARI helps detect, quantify, and control error especially when audio recordings can be matched with response data or screen capture from the interview. In this article, we examine how the use of CARI will reduce or mitigate errors at each step of conducting a survey, and we provide a description of the CARI system currently deployed at the US Census Bureau.

In interviewer-administered surveys, some aspects of total survey error (Biemer 2010) can be managed through use of computer audio-recorded interviewing (CARI) technology. By addressing faulty question wording or translation, poor reading by the interviewer, mode effects, field data entry errors, item nonresponse and other problems, CARI helps detect, quantify, and control error. It is especially helpful when audio recordings can be matched with response data or screen capture from the interview (Thissen et al. 2010).

In this article, we examine how the use of CARI will reduce or mitigate errors at each step of conducting an interviewer-mediated survey, and we provide a description of the CARI system currently deployed at the US Census Bureau. For best value, a complete CARI system consists of:

1. Recording software to collect audio recordings and screen images of the interviewers' and respondents' vocal exchanges during an interview;

2. Transmission, storage and management of the recordings for review;

3. Monitoring system to allow researchers, supervisors, and quality assurance (QA) staff to listen to the recordings at a later time; and

4. Feedback to interviewers and survey designers for quality improvement.

Though this article discusses a specific implementation, similar technology is used in numerous organizations, including governmental data collectors such as Statistics Canada and Statistics New Zealand; academic institutions such as
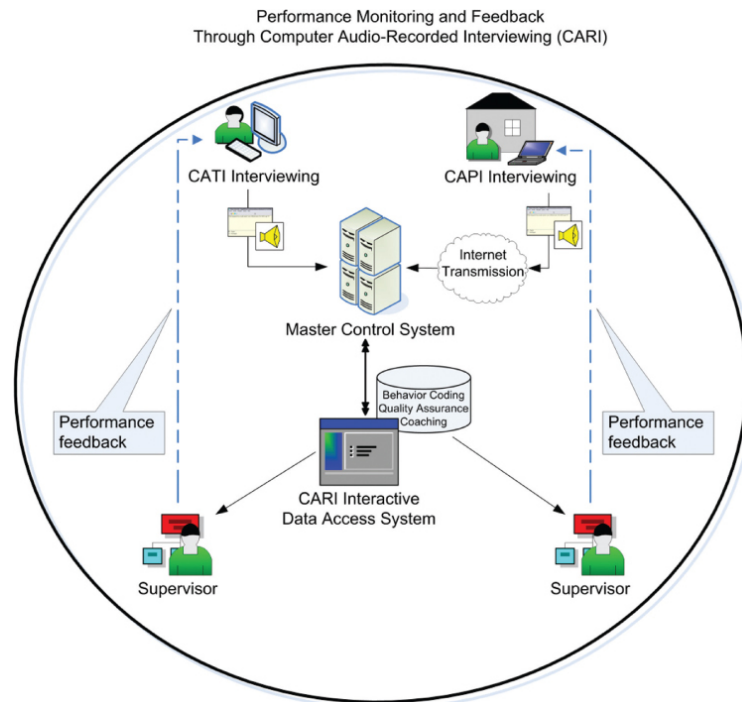
**Figure 1**    Multi-mode performance feedback based on audio recordings (Thissen, Park, and Nguyen 2011).

University of Wisconsin, University of Michigan and others; and commercial or non-profit research firms such as Westat, Inc. and RTI International (for a more complete list, see Thissen et al. 2008).

## CARI Technology and Methodology

The Census Bureau implemented CARI in a series of projects, beginning with feasibility studies in 1999 (Wrenn-Yorker and Thissen 2005) through construction of the current enterprise-wide system (Nguyen, Seige, and Kunta 2012). A simplified diagram of the overall flow of CARI data is shown in Figure 1 for management of multi-mode interviewing and operational performance. Similar workflows allow evaluation of questionnaire design through behavior coding and improvement of data quality through assessment of authenticity, accuracy of data entry, and adherence to interviewing protocol.

In this CARI process, audio recordings and screen images are captured by Blaise software on the computer of a field or telephone interviewer. Regardless of where the data originate, the files move to the centralized CARI Interactive Data Access System, designed by a joint team from the Census Bureau and RTI International. There, audio recordings are played back for coding alongside an image of the question screen with its keyed response data.

At the Census Bureau, specific coding schemes are configured for each survey and type of review. Schemes vary for interviewer coaching, data QA or behavior coding. For example, a coaching coding scheme may include evaluations of interviewers' activities, such as how well they read, tone of voice, or refrain from biasing responses. Interviewers may be complimented when they

**Figure 2** Behavior coding for the ACS Content Test (Pascale 2011, with permission).

|  | CATI | CAPI |
|---|---|---|
| Sample size | 23,673 | 15,202 |
| Completed+partial cases | 4523 | 6384 |
| Recording consent rate | 88.5% | 64.8% |
| Interviews recorded | 4005 | 4137 |
| Interviews selected for coding | 807 | 708 |
| Training cases | 32 | 0 |
| Cases coded out as inaudible | 48 | 7 |
| Interviews coded | 727 | 701 |
| Recordings coded | 27,163 | 22,676 |

overcome difficult situations, persuade reluctant respondents, or maintain control of the environment under challenging circumstances. Similar codes may suit data QA, augmented with accuracy of keyed data. For questionnaire evaluation, respondents' reactions may be quantified through topic-specific coding schemes. The Census Bureau has deliberately built flexibility for addressing multiple aspects of survey error in this way.

The American Community Survey (ACS) Content Test was the first survey to go live with the CARI System at the Census Bureau, employing it in late 2010 and early 2011 (Pascale 2011). Data were collected under two modes: computer-assisted telephone interviewing (CATI) and computer-assisted personal interviewing (CAPI). The behavior coding module of the system was used primarily, to assess new modules and explore improvements to existing ACS questions. The QA module was tested alongside and later with Content Test data. Figure 2 shows details of the Content Test and its use of CARI.

Census Bureau staff who used the CARI System were asked to list the features that they found to be most important. Their comments included the following:

- Analysts cited flexibility in defining coding schemes,
- Coders cited the presence of a screen Image along with the audio recording, offering the exact question wording (including fills) and entered data value,
- Managers and analysts valued support for real-time monitoring of coding quality through inter-rater reliability tests built into the system,
- Supervisors and managers liked having up-to-the-minute data available for extraction.

In general, the field test demonstrated the value of the approach, offering proof of operational value in behavior coding and QA.

## CARI as an Integral Part of Survey Operations

The CARI system supports improved quality and reduction of total error in several ways, being woven integrally into preparation and execution of survey data collection processes. Consider the basic steps taken for any survey, defined simply below, and how the use of audio recording and review might alter the operational approach.

- **Step 1. Define Research Objectives and Population.** Audio recording cannot assist in selecting the target population and preparing the sample, and so those actions are outside the scope of this discussion. However, when defining research objectives, certain items in each survey are key analytical data points, such as demographic and lifetime-smoking questions in a survey of tobacco use. These survey questions can be targeted for audio recording to reduce the error level of responses that are critical for analysis.

- **Step 2. Choose Methods of Data Collection.** Survey results may be affected by the mode of data collection. For example, respondents are less likely to report honestly to sensitive questions through in-person interviewing relative to computerized data collection (Villarroel et al. 2008). Another example is the recency effect, whereby respondents select the most recent item given in response options for the question. Elderly respondents, for example, tend to endorse the latest answer option more often in telephone surveys compared to surveys where answer options are provided visually, due to reduced ability to remember spoken options (Krosnick and Alwin 1987).
  If multiple electronic modes are anticipated, surveys can attempt to identify and potentially reduce the level of error. Specific questions suspected of exhibiting mode effects, such as those with long lists of response options, can be targeted for recording in all modes. Following data collection, audio recordings can be coded for characteristics of respondent behavior, enabling comparison of differences across modes. In this example, if respondents more frequently ask for repetition of the response options during telephone interviews compared to the same item during in-person interviews, one would infer that the option lists were too lengthy to be presented over the phone although acceptable for face-to-face interviewing.

- **Step 3. Plan for Quality.** QA plans profit from CARI as a tool for detecting, remediating, and controlling various types of error. Audio recordings provide abundant data by capturing "live" interviewer-respondent interactions and can provide a verifiable reference for the responses that should have been entered. Authenticity of responses may be hard to prove, but falsification sometimes can be detected easily through CARI, such as when the interviewer keys in data at

home without contacting the respondent; in that case, the audio recording will capture the sound of keystrokes but not the verbal interchange expected between interviewer and respondent. When the quality plan includes CARI review to confirm authenticity, curbstoning can be detected and addressed.

- **Step 4. Construct and Pretest the Questionnaire.** Items with awkward, lengthy, or complex wording, especially among the response options, may increase error for some respondents, due to cognitive fatigue, impatience, or failure to understand, and those problems can be detected by CARI (Mitchell et al. 2008). Omitting a specific reference period for temporal questions can cause unexpected errors as can region-specific vocabulary such as "bubbler" for water fountain or drinking fountain (Remlinger, Salmons, and Schneidemesser 2009) or cultural expectations (which days are holidays). The omission of the word "usually" changes the meaning of "By what form of transportation did you usually travel to work?" from a long-term description to one interpreted as referring to the present day only. In multilingual surveys, translation error may not deliver the intended message consistently. CARI recordings provide evidence of situations in which the respondent becomes tired, confused or misled by the phrasing of the question. In a pretest environment, this approach can provide valuable insight for improving a questionnaire, offering a great opportunity to reduce error.

- **Step 5. Collect Data.** It is a challenge to detect and quantify errors of data collection. Audio recording and review in CATI or CAPI mode allows monitors to assess interviewer performance and data quality. For example, interactions between interviewers and respondents can affect responses, causing bias in the resulting data or affecting item level nonresponse rates. Though this potential source of bias has long been recognized and remains an important concern today, most methods for direct evaluation of interviewer-respondent interactions, such as live observation in the field, may themselves introduce bias into the interactions. Silent monitoring in a call center does not introduce bias, but it adds to staff workload, requiring monitors to listen to non-productive time such as dialing, busy signals and answering machines as well as live interviewing. Audio recording allows observation without having an impact on the interviewing environment and without wasting time observing non-interview activities.

  During data collection, data entry problems may occur when interviewers key lengthy open-ended responses. In one study, audio

recordings replaced field data entry entirely after it was determined that the quality of field data entry was unacceptable compared to transcription from recordings (Edwards, Hicks, and Carlson 2010).

- **Step 6. Analyze and Report.** Typically, CARI has not been used to augment reports of findings. However, recordings could be provided in a multi-media-format report, or they could be quoted as case-based evidence of unusual findings.

## Conclusions

CARI technology offers a means of controlling levels of survey error through methods that have become well established or are just beginning to be explored (Hicks et al. 2010). Already its value has been shown for questionnaire evaluation, interviewer performance management, detection of falsification, and assessment of response data quality. Adoption of the technology is spreading, as organizations recognize that the startup efforts of defining and implementing a new workflow pay back in quality and expectations of cost savings. In the future, audio review information as a form of paradata may be linked to other paradata, such as interviewer or respondent characteristics, or to other operational information, to give a wider view of error sources of a survey and how to remediate them.

## Acknowledgements

# REFERENCES

Biemer, P.P. 2010. "Total Survey Error: Design, Implementation and Evaluation." *Public Opinion Quarterly* 74 (5): 817–48. http://poq.oxfordjournals.org/content/74/5/817.abstract.

Edwards, B., W. Hicks, and M. Carlson. 2010. "Old Dogs New Tricks: CARI and CASI Innovations to Reduce Measurement Error and Nonresponse." In *The International Total Survey Error Workshop*. Stowe, VT. http://www.niss.org/events/itsew-2010](http://www.niss.org/events/itsew-2010.

Hicks, W.D., B. Edwards, K. Tourangeau, B. McBride, L.D. Harris-Kojetin, and A.J. Moss. 2010. "Using CARI Tools to Understand Measurement Error." *Public Opinion Quarterly* 74 (5): 985–1003. http://www.oxfordjournals.org/our_journals/poq/special.html.

Krosnick, J.A., and D. Alwin. 1987. "An Evaluation of a Cognitive Theory of Response-Order Effects in Survey Measurement." *Public Opinion Quarterly* 51 (2): 201–19. http://comm.stanford.edu/faculty/krosnick/docs/1987/1987%20an%20evaluation%20of%20a%20cognitive%20theory%20of%20response-order%20effect.pdf.

Mitchell, S.B., M.M. Strobl, K.M. Fahrney, M.T. Nguyen, B.S. Bibb, M.R. Thissen, and W.I. Stephenson. 2008. "Using Computer Audio-Recorded Interviewing to Assess Interviewer Coding Error." In *Proceedings of the Survey Research Section, American Statistical Association*, 4414–21. http://www.amstat.org/sections/srms/proceedings/y2008/Files/strobl.pdf.

Nguyen, M., B.C. Seige, and N. Kunta. 2012. "CARI Interactive Data Access System: A Robust and Scalable System for the Review of CARI Recordings." In *The International Field Directors and Technologies Conference*. Orlando, FL. http://www.ifdtc.org/PC2012/presentation_2012_files/9E%20CARI%20Interactive%20Nguyen.pdf.

Pascale, J. 2011. "Using Behavior Coding to Evaluate Questionnaires." In *The Federal Computer Assisted Survey Information Collection. Workshop*. Washington, DC. https://fedcasic.dsd.census.gov/fc2011/ppt/03_pascale.pdf.

Remlinger, K., J. Salmons, and L.Von Schneidemesser. 2009. "Revised Perceptions: Changing Dialect Perceptions in Wisconsin and Michigan's Upper Peninsula." *American Speech* 84 (2): 176–91. http://americanspeech.dukejournals.org/content/80/1/22.short.

Thissen, M.R., C. Fisher, L. Barber, and S. Sattaluri. 2008. "Computer Audio-Recorded Interviewing (CARI), a Tool for Monitoring Field Interviewers and Improving Field Data Collection." In *The International Methodology Symposium*. Gatineau, Canada: Statistics Canada. http://www.statcan.gc.ca/pub/11-522-x/2008000/article/10955-eng.pdf.

Thissen, M.R., H. Park, and M. Nguyen. 2011. "Computer Audio Recording: A Practical Technology for Managing Survey Quality." In *The International Total Survey Error Workshop*. Quebec City, Canada.

Thissen, M.R., S. Thorpe, E. Peytcheva, C. Barnes, H. Park, and C. Fisher. 2010. "Improving Data Quality through Audio and Image Recording." In *The International Total Survey Error Workshop*. Stowe, VT. http://www.niss.org/events/itsew-2010](http://www.niss.org/events/itsew-2010.

Villarroel, M.A., C.F. Turner, S.M. Rogers, A.M. Roman, P.C. Cooley, A.B. Steinbeg, E. Eggleston, and J. Chromy. 2008. "T-ACASI Reduces Bias in STD Measurements: The National STD and Behavior Measurement Experiment." *Sexually Transmitted Diseases* 35 (5): 499–506. http://www.ncbi.nlm.nih.gov/pubmed/18434943.

Wrenn-Yorker, C., and M.R. Thissen. 2005. "Computer Audio Recorded Interviewing (CARI." In *Technology. Presented at the Federal Computer Assisted Survey Information Collection Workshop*. Washington, DC. https://fedcasic.dsd.census.gov/fc2005/index.php.