



Implementation of CARI Technology to Improve Data Quality

Hyunjoo Park
Rita Thissen
Paul P. Biemer

International Workshop
on Internet Survey and Survey Methodology

Introduction

- Prevalent use of surveys has adversely affected response rates and potential respondents' (Rs) impressions about surveys.
- The widespread adoption of technology, especially the use of computers for surveys, pushes many aspects of survey data collection in positive directions.
- Digital audio recording is one promising component with great potential for improving survey quality and reducing the burden on fatigued respondents.

Computer Audio-Recorded Interviewing (CARI)

- CARI is a technology that allows the computer to act as a sound recorder while the interviewer (IWR) administers a computerized survey questionnaire.
- CARI records the verbal exchange between IWR and R.
- The recording can be switched on or off by the computer for all interviews, for portions of interviews, or for randomly selected interviews.
- This paper reviews how CARI technology can be used for many aspects of surveys in order to improve data quality.

The Cycle of CARI

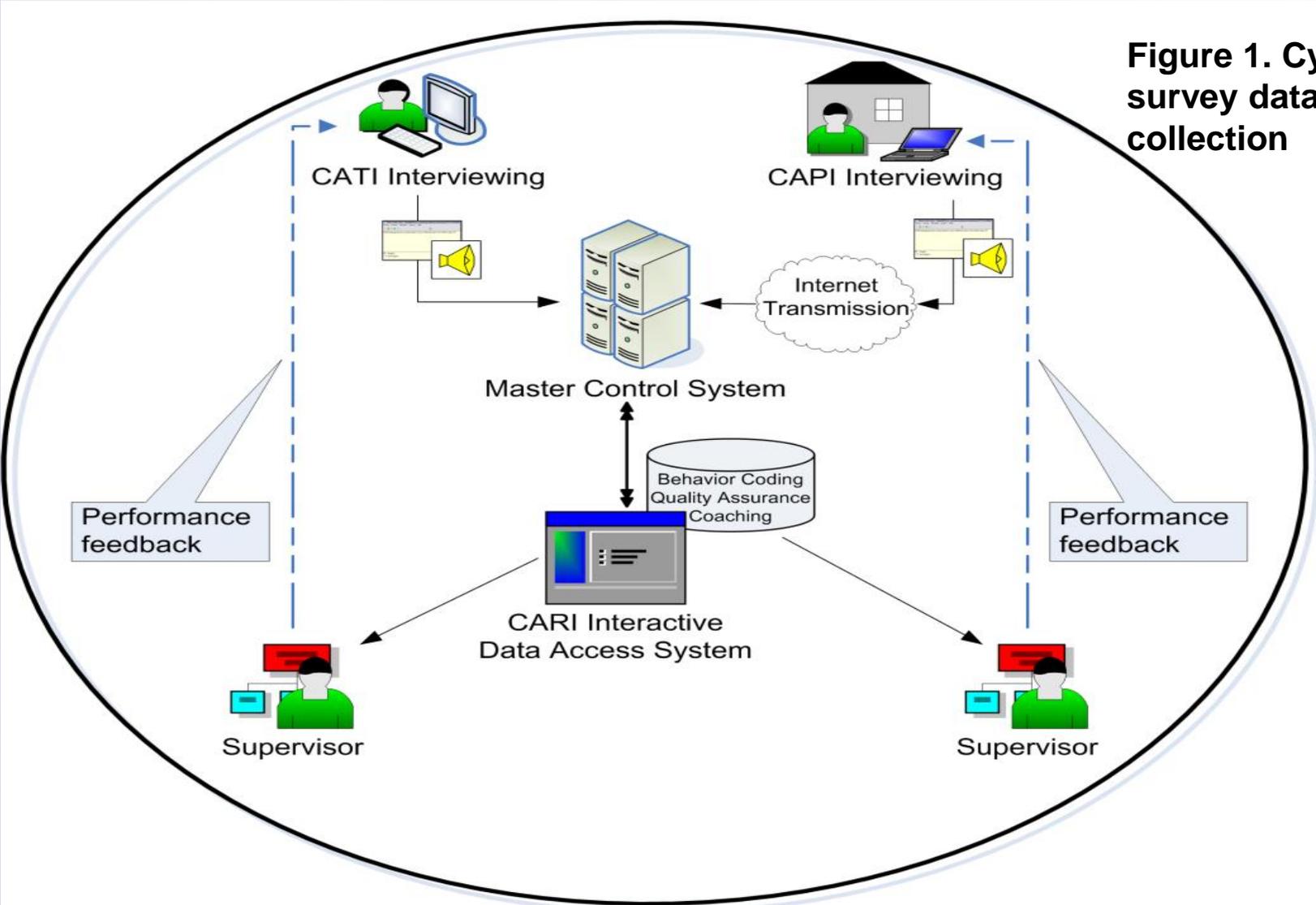


Figure 1. Cycle of survey data collection

The Cycle of CARI: Step 0. Startup Activities

- Several startup activities must take place prior to the beginning of data collection including:
 - system developers create or configure infrastructure to support data collection;
 - survey specialists design and pre-test the questionnaire;
 - and statisticians select the sample.
- It should be noted that the decision on whether to use CARI technology will affect these startup activities.

The Cycle of CARI: Step 0. Startup Activities – *Cont'd*

Examples:

- 1) Adequate time should be allocated at the questionnaire development stage because CARI implementation requires additional time for developing and testing the programmed questionnaire (instrument).
- 2) When CARI is to be implemented, the configuration files should be modified so that it can activate and de-activate the microphone at specific items or times.
- 3) The instrument may also initiate and terminate calendar entries, global positioning system data capture, video recordings, or screenshots.

The Cycle of CARI: Step 1. Interviewing and Transfer

Forms Answer Navigate Options Help

Where did you spend your childhood years, from birth to age 14?

1. All or mostly in the USA
 2. All or mostly outside the USA
 3. A mixture of US and non-US locations

Figure 2. Example of a simple questionnaire item programmed with Blaise

The Cycle of CARI: Step 1. Interviewing and Transfer – *Cont'd*

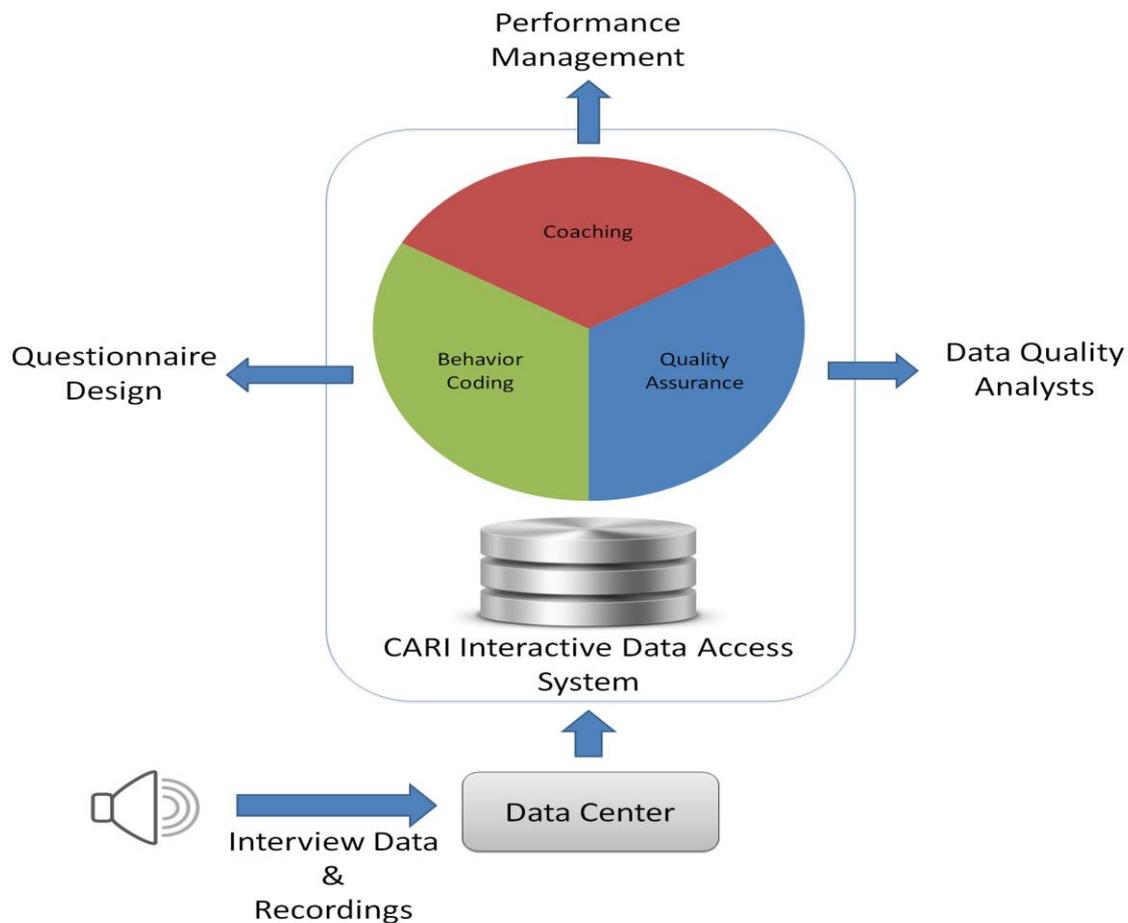


Figure 3.
Data flow and systems that support CARI data collection, storage, and review

The Cycle of CARI: Step 2. Data Storage

- Data storage requirements depend on characteristics of the survey:
 - instrumentation software,
 - system(s) employed for review,
 - size of the survey's respondent pool, and
 - size of the response dataset.

- The chosen method typically meshes with the data transfer system for receipt and storage on a flow basis, allowing quick turnaround of review and feedback.

The Cycle of CARI: Step 3. *Quality & Operational Review*

- Data stored at the central site are made available to quality review staff, supervisors, and others through some type of CARI review system, which may be
 - as simple as audio playback software with a mechanism for keeping notes.
 - as complex as a large commercial relational database with a role-based web application for playback, coding, and tracking of the cases and their audio segments, along with functionality for management of the review operation itself, as is found at the US Census Bureau.
 - It may fall in between, as does the one at RTI International, with a single coding interface and versatile scoring.

The Cycle of CARI:

Step 3. *Quality & Operational Review – Cont'd*

- Review may include all/a subset of response files, audio recordings, screen capture, video recording, GPS coordinates, calendar entries, document scans, assessment scores, or any other information collected during the interview.
- Review results pass to supervisors who recognize accomplishments or address concerns raised from review of recordings combined with other information.

CARI
Interactive Data Access

Log Out

Workload

View All Questions

- + CAPI100300
- + CAPI100301
- + CAPI100302
- CAPI100303
 - [1] - GENDER_IN...
 - [1] - GENDER_CLI...
 - [1] - AGE
 - [1] - AGE_CLIEN...
 - [1] - PITCH_INT...
 - [1] - PITCH_CLI...
 - [1] - BIRTH_COU...
 - [1] - BIRTH_COU...
 - [1] - LANGUAGE_...
 - [1] - LOUDNESS_...
 - [1] - LOUDNESS_...
 - [1] - SPEED_INT...
 - [1] - SPEED_CLI...
 - [1] - ROOMNOISE...
 - [1] - ROOMNOISE...
 - [1] - COMMENT
- CAPI100305
 - [1] - SPEED_CLI...
 - [1] - SPEED_CLI...
- + CAPI100307
- + MIX_100321
- + MIX_100322
- + MIX_100323

User: sysadmin Component: QAC
Role: System Administrator Survey Name: Survey for testing QAC CATI Excl. Survey Mode: Multi-Mod

Blaise Data Entry - \\rtpwfil02\cai_system

Forms Answer Navigate Options Help

📎 ? ! 🗣️

To categorize the voice characteristics of the a

For example, my gender is
(INTERVIEWER: SELECT THE APPROPRIATI

Question	Recording Date Time
GENDER_INTERVIEWER	1/10/2010 12:43:47 PM
GENDER_CLIENT	1/10/2010 12:44:54 PM
AGE	1/10/2010 12:45:00 PM
AGE_CLIENT	1/10/2010 12:45:10 PM
PITCH_INTERVIEWER	1/10/2010 12:45:33 PM
PITCH_CLIENT	1/10/2010 12:45:49 PM
BIRTH_COUNTRY_INT	1/10/2010 12:46:02 PM
BIRTH_COUNTRY	1/10/2010 12:46:21 PM

Categories	Speed, not too slow/fast	Volume, not too loud/soft	Clear pronunciation
Authenticity	S4 Strongly agree <input checked="" type="radio"/>	L4 Strongly agree <input type="radio"/>	P4 Strongly agree <input checked="" type="radio"/>
Reading	S3 Somewhat agree <input type="radio"/>	L3 Somewhat agree <input checked="" type="radio"/>	P3 Somewhat agree <input type="radio"/>
Interview administration	S2 Somewhat disagree <input type="radio"/>	L2 Somewhat disagree <input type="radio"/>	P2 Somewhat disagree <input type="radio"/>
Probing	S1 Strongly disagree <input type="radio"/>	L1 Strongly disagree <input type="radio"/>	P1 Strongly disagree <input type="radio"/>
Conduct			
Other			

Codes:

Notes:

 Needs Secondary Review Needs Questionnaire Review

Category	1	2	3	4	5	6	7	8	9	10	Notes
Authenticity	V5	G1									Edit
Interview administration	S4	L3	P4								Update Cancel Delete
Conduct	M4	B3	F4								Edit

Disclosure Prohibited - Title 13 U.S.C.

12

CARI to Improve Data Quality

- Proven capabilities including:
 - monitoring field interview staff,
 - confirming interview authenticity & detecting curb-stoning,
 - diagnosing questionnaire problems,
 - offering an intimate view of in-person interviews & performance information
 - improving survey data quality by replacing/supplementing field verbatim transcription of responses.
- CARI might offer a remedy to resolve declining response rate issues through the analysis of conversation between interviewer and respondent.

- 1. Reducing IWR related survey errors***
- 2. Reducing R related survey errors***
- 3. Reducing other survey errors***

Survey Errors Related to Interviewers : Role Dependent Errors

- ***Role-independent*** refers to cases where respondents' answers are affected by interviewer presence or their demographic characteristics.
 - i.e. Rs' answer to the race related questions vary depending on the interviewer race.
- This type of error is known to appear for certain topics of surveys, so deliberate interviewer-respondent match, in terms of demographics or, for example, appropriate interviewers' attire for the sampled neighborhood, can resolve most of these role-independent interviewer effects.

Survey Errors Related to Interviewers : Role Dependent Errors – Cont'd

- ***Role-dependent error*** is caused by interviewers who break the rules of the standardized interview.
- This happens regardless of the topic despite researchers' effort via intensive interviewer training.
 - Novice interviewers may commit this error in ignorance of the basic rules of standardized interviews while veteran interviewers through bad habits matured along their career as a survey interviewer.
- CARI's unobtrusive characteristics of data recording enable us to access this information while decreasing the costs associated with field observation.

Survey Errors Related to Interviewers : Short-cutting & Data Fabrication

- Confronted by the historically lowest response rates, Interviewers may be tempted to fabricate the interview.
 - Fictitious interview
 - Curb-stoning. i.e., skipping lengthy sections by falsifying the response to gateway questions that lead to a subsequent series of more detailed questions.
- CARI reviewer can confirm validity by noting the behaviors on key questions, listening for two distinct voices, and confirming reasonable response patterns.
- Specific circumstances that are suspicious can be flagged for more intensive review.

Survey Errors Related to Interviewers : Interviewing Performance Monitoring

- A great management and training tool for field managers since each recorded segment., provides direct evidence of performance and data quality.
- Survey managers configure scoring algorithms based on the codes that are defined for the specific survey.
- By using this system, survey managers can easily identify interviewers who need further training and obtain knowledge about the areas needing improvement for each interviewer.

Staff Workload

Coaching Compor

Enter search values and click the Filter button.

Status: Current Code: Intv Site:
 Intv Date From: Intv Mode: Intv ID:
 Intv Date To: Language: Control Num:
 SSF: FS:

Select	Mult Site	Mult Intv	IRR	Type	Control Num	Status Date	Current Code	SSF	FS	Coding Site	Resp Language	Treat	QA Auth	QA Rec Qual	QA Data Qual	Coach Rec Qual
<input type="checkbox"/>				CC	CNTR600600	12/31/2012	bssr_rj	ssfrd002	fs02rd02	Richmond	English	0	3	3.3	83.3	
<input type="checkbox"/>			🟢	CC	CNTR600600	12/27/2012	demo_rm1	ssfrd002	fs02rd02	Richmond	English	0				
<input type="checkbox"/>			🟢	CC	CNTR600600	12/27/2012	cdr_rj	ssfrd002	fs02rd02	New York	English	0				
<input type="checkbox"/>			⬆️⬆️	CC	CNTR600601	12/27/2012	demo_rm1	ssfrd001	fs01rd01	Richmond	English	1	1.7	3		5
<input type="checkbox"/>			🟢	CC	CNTR600601	12/27/2012	cdr_rj	ssfrd001	fs01rd01	New York	English	1				
<input type="checkbox"/>				CC	CNTR600602	12/31/2012	bssr_rj	ssfrd002	fs02rd02	Richmond	English	0	3.3	3.5		5

- 1. Reducing IWR related survey errors***
- 2. Reducing R related survey errors***
- 3. Reducing other survey errors***

Survey Errors Related to Respondents

- Rs make errors intentionally or unconsciously and the level of errors can be decreased or increased with the presence of interviewers.
- The errors initiated from Rs occur along the cognitive process consist of comprehension, information retrieval, formulating an answer, and formatting and editing responses (Bradburn 2004)

Survey Errors Related to Respondents : Comprehension

Cognitive Process	Definition and Examples
Comprehension	<p>The goal for the researcher is for Rs to understand the question in the same way that the researcher does.</p> <p><i>i.e., misunderstanding in questions using common terms such as “weekday”, “children,” and “regularly.” (Belson 1981)</i></p>

Survey Errors Related to Respondents : Information Retrieval

Cognitive Process	Definition and Examples
Information Retrieval	<p>It is process by which the memory storehouse is searched to retrieve a particular item sought. It is facilitated by cues in the question that activate the pathways of association leading to the desired information.</p> <p>i.e., Telescoping</p>

Survey Errors Related to Respondents : Formulating an Answer

Cognitive process	Definition and Examples
Formulating an Answer	<p>Rs must formulate an answer to the question. If the questions are about behavior that is not well remembered or about attitudes that have not been well discussed, Rs construct answers on the spot using all the information from available sources in working memory.</p> <p><i>i.e. Assimilation, Contrast</i></p>

Survey Errors Related to Respondents : Formatting and Editing Responses

Cognitive process	Definition and Examples
Formatting and Editing Responses	<p>Rs must find their answers among the response options that IWR offers. Even after a response option has been chosen in Rs' mind, they may still edit the response out of their concerns with self-presentation.</p> <p><i>i.e., Primacy, Recency, Social Desirability,</i></p>

How CARI Detect Respondent Related Errors

- The audio recordings help researchers detect many of the respondent originating errors and shed light on how to revise questionnaires to reduce such errors.
- When a situation conducive to error is suspected, the questionnaire may be modified or interviewers can be trained to provide better guidance.

How CARI Detect Respondent Related Errors – Cont'd

- Examples:
 - Identify problematic terms by detecting accidental errors that occur when R comprehends the question incorrectly.
 - Identify problematic words as IWR may stumble or R may ask to read or hear the question again.
 - Identify too lengthy question text or list of response options as IWR may paraphrase, or R may volunteer to answer the question before listening to all response options.
 - Identify problematic wording by watching Rs' uncomfortable reactions when a sensitive question is asked or perceiving the nuances associated with the selection of certain response options.

- 1. Reducing IWR related survey errors***
- 2. Reducing R related survey errors***
- 3. Reducing other survey errors***

Reducing Other Survey Error

- IWRs' code answers on the spot using the pre-existing categories; however, it can be improved (i.e. I/O coding)
- Audio-recordings resolve this issue by bringing the appropriate context associated with respondent open-ended answers to the appropriate coding experts.
- Correction of errors in the data entry step for a discrepancy between the entered response and the respondent's reporting of the same survey question.

Conclusion

- The use of CARI digital audio recording technology puts greater knowledge and power into the hands of data collection managers.
- Though implementation may vary, each operation can benefit from the cycle of data collection, review, and feedback.
- The potential of CARI technology in survey research remains very promising and worthy of further investigation.

More Information

Hyunjoo Park

Research Survey Methodologist

415-848-1393

mpark@rti.org