

Paradata and Blaise

A Review of Recent Applications and Research

Jim O'Reilly
Westat, USA

Agenda

- Evolution of Paradata
- Significant Recent Extensions and Application
 - Census Bureau and PANDA
 - Statistics Canada paradata strategy
 - Statistics Canada application—POINT
 - National Health Interview Survey—public use paradata and impetus to extend paradata
- CARI – paradata’s Next Generation
 - Census Bureau field test
 - 2008 American National Election Survey
 - Westat’s integrated CARI system

Evolution of Paradata

- Modest beginnings in 1980s as trace files
 - Fields entered, keys pressed, timestamp
 - Stream of entries with little structure
 - Used for debugging, recovery of lost data, testing
- Blaise audit trails implemented in Blaise III (?)
 - More structure
 - Modular DLL implementation
 - Extensible—key strokes, pointer coordinates, custom functions – CARI, etc.
- Improvements fostered wider application

Evolution of Paradata

- Use varies widely
- Some use paradata rarely or not at all
 - Overhead to implement
 - Legacy systems meet needs
 - Value seen as limited—”just methodology”
 - Customers not aware/not convinced of benefits
- Others use audit trails soberly
 - Summary reports on field/block timing stats for management
 - Interviewer statistics on field/block timing to spot outliers
 - Archiving adt’s for ad hoc uses
 - data recovery, special investigations
- Others view paradata as a critical tool for survey quality
 - Apply it comprehensively
- Let’s review cutting edge work crossing threshold to a new standard

Census Bureau and PANDA

Ari Teichman (2009)

- Performance and Data Analysis (PANDA) system
- Implemented in 2007 American Housing Survey
- Goal: improve survey quality by providing early warnings of
 - Interviewers having difficulty with key survey concepts
 - Possible falsification
- Paradata-based reports to managers on key metrics
 - Interview duration, ivw by time of day, ivw result, outliers
 - Possible falsification
 - high rates of vacant housing, small hh size, ivws at unusual time, short interviews

Census and Panda

- Management reports
 - Summary level on regional office/area totals, cumulative/weekly report, average cases per interviewer and interviewer reports on
 - Highest non-response for salary of R
 - Highest regular ivws completed in <20min
 - Highest # of case completed 12:00am – 7:59am
 - Field managers can drill down in ATs files to study details; FM's said
“they used the system to search for detailed information on interviewers' work, to address potential problems appropriately, and to identify interviewers retraining or cases requiring re-interviews.”
- System well accepted by staff
- Being implemented in other major Census surveys, beginning with the National Health Interview survey.

Statistics Canada Strategic Paradata Approach

François LaFlamme (2009)

- StatCan has made a major commitment “to data collection research using paradata as the cornerstone”
- Goals: understand the process, develop new efficiencies, evaluate new initiatives, and maintain and improve data quality
- Data collection a top concern: determines data quality and accounts for 50-75% of total survey costs
- Developed a paradata warehouse storing call and contact information for tel and in-person interviews, admin and payroll data
 - Centralized warehouse cuts burden on studies and customers insuring all surveys are represented
- System provides in-depth, timely cost data for survey cost analysis

Statistics Canada Strategic Paradata Approach

- Early research efforts focused on CATI
 - Analyzed time of contact attempts and system work, contact rates, calling patterns and production-cost relationship
 - Found capping # of calls reduced survey cost by 3.1 to 4.2% for longitudinal surveys
- StatCan expect improvements from system in:
 - Better use or pre-collection data and data gathered during collection
 - Methods used after first contact
 - Development of a responsive design framework
 - Predicting collection requirements during collection based on progress metrics.

Statistics Canada – POINT

Mike Maydan (2009)

- On management challenges in complex survey context
 - Regional collection structure, multiple archives, reporting systems and competing priorities, needs and dimensions
- To improve coordination and integration developed
 - Reports on data accuracy with response/non-response rates, non-response follow-up, refusal conversion and tracing
 - Based on case-level paradata from CATI history file and CAPI case management events
 - Pace of Interview (POINT) system focused on irregular production calls
 - Based on audit trails; evaluating the “act of collection”

Statistics Canada – POINT

Mike Maydan (2009)

- POINT designed to apply objective performance measures to help identify interviewers for possible retraining or other action. Based on
 - Pace of the interview (field changes per minute)
 - Item non-response (don't know, refusal)
- Threshold for suspect irregular call levels derived from early collection period—600 calls with ≥ 20 changed fields. Threshold set at 175% of early collection mean field changes per minute, and $>25\%$ item non-response
- Detailed daily report provided to managers

Public Use Paradata and Research

Beth L. Taylor (2009)

- National Center for Health Statistics includes paradata on the data collection process along with standard public data file
 - 2006 National Health Interview Survey
 - Annual survey of 35,000 families--in-person & tel follow-up. Detailed contact history information on each attempt: description, reluctance encountered, strategies to complete. For non-contacts description and strategies kept
 - PD includes interview language, cooperativeness of respondent, interview mode, reasons for interview breakoffs, type of non-interview cases, time of interview and module/section times.
 - From audit trail: time per question, dates, and interviewer notes.
 - Recodes in public data against individual id

Public Use Paradata and Research

Beth L. Taylor (2009)

- Possible analyses
 - Contact attempts and interview completion
 - Time of day of interview
 - Interview strategies and successful completion
 - Characteristics of hard-to-contact families
 - Modeling impact of interview mode on health outcome
- 2008 NHIS data release will add PD for visit attempts, use of function keys and language of interview
- Working with Census to enhance interviewer performance, tracking, and reporting in PANDA



Public Use Paradata and Research

Beth L. Taylor (2009)

- NHIS staff studying interviewer performance, using case level contact histories and audit trail item and section
- Other ongoing research on
 - Process of selecting cases for reinterview and applying statistical predictors for re-interviews
 - Non-response adjustment to weights
 - Sub-unit response
 - High-effort interviews and bias

Other paradata research to be presented at 2009 Joint Statistical Meetings

- “Use of Paradata to Manage a Field Data Collection”, Robert Groves (University of Michigan), et al.,
- “Using the Fraction of Missing Information to Monitor the Quality of Survey Data”, James Wagner (University of Michigan)
- “Modeling the Difference in Interview Characteristics for Different Respondents”, John Dixon (Bureau of Labor Statistics)
- “An Evaluation of Nonresponse Bias Using Paradata from a Health Survey”, Aaron Maitland (National Center for Health Statistics) et al.
- “Subunit Nonresponse in the National Health Interview Survey (NHIS): An Exploration Using Paradata”, James M. Dahhamer National Center for Health Statistics and Catherine M. Simile (NCHS)

Summary

- Paradata applications have advanced and matured
- Organizations integrating paradata into core management process
- Providing carefully developed metrics and reports to various supervisory levels
- Giving direct access to detailed audit trail information for line supervisors.
- Further research blooming

CARI -- Paradata's Next Generation

- CARI
 - Audio recording of the interviewer and respondent during the interview
 - Enabling review and analysis of a multitude of facets of the interaction, far beyond that of audit trails
 - Used most for QA focused on detecting falsification and evaluation of interviewer performance
 - As with PD, seems on cusp of shift from exploratory and specialized application to general use and comprehensive scope

Census Bureau CARI Field Test Evaluation

Arceneaux (2007)

- Part of broader effort toward using “CARI in all of the Census Bureau’s computer-assisted personal interview (CAPI) surveys.”
- 2006 study of 423 recorded cases in three regions
- Found CARI
 - functioned properly, recording occurred without detection, and technical problems did not increase
 - Rs were very receptive to CARI while interviewers were mixed-- 39% comfortable and 23% opposed.
- Two negative findings
 - Recordings were rated excellent or good for 85.6% of cases, while the desired level was 96%.
 - Survey response rate was 81% compared to 90% on a comparable sample from the NHIS

Census Bureau CARI Field Test Evaluation

Arceneaux (2007)

- Arceneaux mentions a number of factors in the field test study that may mitigate these differences and recommends further research.

2008 American National Election Survey

Lupia, et al. (2009)

- CARI used to verify that pre-election and post-election interviews completed
- Cases reviewed based on address validation, respondent self-identification, confirmation of both respondent and interviewer voices, and consistency of voice(s) across the recordings
- Any concerns prompted review /in-person field validation
- Percentage of cases reviewed
 - 100% during the initial interviewer incentive period
 - Later at least 10% of completed cases per interviewer

A System Approach for Using CARI

Wendy Hicks et al. (2009)

- Full system approach to CARI to meet varied research objectives.
- In central office CARI recordings are linked to the survey data and audit trails.
- Flexible review in different modes
 - Coders can review one question across a series of interviews, or a series of questions in an individual interview.
 - System provides coding with a structured tool
 - Links to the survey data and operational process data
 - Other built-in quality control functions
 - A variety of reports.

A System Approach for Using CARI

Wendy Hicks et al. (2009)

- Application evaluating questions and interviewer performance in a national health survey
- Examined
 - Whether respondent asked for clarification
 - Whether response did not match the format expected
 - Interviewer skills and the percentage of interviewers with difficulty in terms of
 - minor changes from verbatim
 - major changes from verbatim
 - not probing when needed
 - using leading probes

A System Approach for Using CARI

Wendy Hicks et al. (2009)

- Combination of a comprehensive and effective system to integrate the management, processing, and coding, along with valuable analytic findings, appears to interest many in the potential of CARI

Two Other CARI Items

NIH BIG Data Initiative for surveys of the future

Mabry and Philogene (2009)

- For surveys of the future CARI recognized as one of four notable developments.

Blaise implementation of CARI in the standard system

- Primary survey interviewing system for complex surveys will have CARI built-in
- Supporting highly flexible and dynamic CARI methods and functions.

Final Thoughts

- Evolution of paradata in survey research seems to be reaching an important threshold.
- Technology, organizational systems to integrate key processes, and research demonstrating the value of paradata to better understand important and challenging survey elements all seem to be aligning.
- Reaching stage for much wider adoption and utilization by the largest organizations, as well others

References

- Arceneaux, Taniecea A.. 2007. "Evaluating the Computer Audio-Recorded Interviewing (CARI): Household Wellness Study (HWS) Field Test". Proceedings of the Survey Research Methods Section, American Statistical Association
- Hicks, Wendy., Brad Edwards, Karen Tourangeau, Laura Branden, Drew Kistler, Brett McBride, Lauren Harris-Kojetin and Abigail Moss. 2009. "A System Approach for Using CARI in Pretesting, Evaluation and Training" *FedCasic Conference, Washington DC*
- LaFlamme, François. (2009) "Overview of CATI Data Collection Research Focused on Developing Operational Strategies for Process Improvement". *FedCasic Conference, Washington DC*
- Lupia, Arthur, Jon A. Krosnick, Pat Luevano, Matthew DeBell, and Darrell Donakowski. 2009. *User's Guide to the Advance Release of the ANES 2008 Time Series Study*. Ann Arbor, MI and Palo Alto, CA: the University of Michigan and Stanford University.
- Mabry, Patricia L. and G. Stephane Philogene, 2009. "Systems Science Methodologies To Protect and Improve Public Health". *IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology, Nashville TN.*
- Maydan, Michael J. (2009) "Using Paradata to Monitor Survey Quality in Statistics Canada's Regional Data Collection MIS Reports". *FedCasic Conference, Washington DC.*
- Taylor, Beth.L. 2009 "The 2006 National Health Interview Survey (NHIS) Paradata File: Overview and Application". *FedCasic Conference, Washington DC.*
- Teichman, Ari. 2009. "Panda: Using Paradata to Improve Data Quality". *FedCasic Conference, Washington DC*