Paradata and Dashboards in PIAAC
Leyla Mohadjer and Brad Edwards, Westat

Managing Data Quality in Large Scale Assessments
May 11-12
OECD
Paris, France
Overview

- Introduction to performance dashboards
- Evolution of paradata discovery dashboard at Westat
- Case study: U.S. implementation of first cycle of PIAAC, with a focus on detecting fabrication
- Future directions
- Summary and conclusions
Introduction to Performance Dashboards
Why Use a Dashboard?

- Enables the driver to keep moving while checking critical systems
- Dashboards decrease risk, increase efficiency
- Surveys can benefit from dashboards in many ways
- Survey operations in the PIAAC countries move very fast, run many risks in production, costs, and quality
- Performance dashboards can help the survey “drivers” monitor how they are doing and signal when something may be going off course
What Is a Dashboard?

A dashboard is a...

visual display of the

most important information needed to achieve one or more objectives; consolidated on a single screen so the information can be monitored at a glance.

Stephen Few (2013)
Parsing the Definition

- A visual display
  - Expertise required to visualize information so the user can process it quickly and accurately; pre-attentive processing

- The most important information
  - User input required to help define it
  - Easy to err by providing too much information
  - Metrics are drivers, change agents to meet objectives

- On a single screen
  - Working memory can only hold 3 or 4 “objects” at a time
Business Dashboards

- Information explosion
- A tool, a communication medium, to control information
- Dashboards made their first appearance in business organizations in the 1990s
- For every good business dashboard, 1000s of bad
  - Dense array of data
  - Small screen real estate
  - Must leverage power of visual perception
Who Is the User?

- If the organization and its IT infrastructure is ready, the most important first step is determining the user.
- Example: field supervisor on a household-based survey.
- Best practices
  - Focus on one user type
  - Identify most important information to them
  - Prototype, test, iterate to incorporate feedback, enhance utility and user acceptance.
Metrics and Key Performance Indicators (KPIs)

- Metrics are change agents
- KPIs are metrics that are directly tied to the overall objectives
- For household surveys in the field, two KPIs stand out:
  - Response rate
  - Hours (or cost) per completed interview
- Examples of other useful metrics at the survey level
  - Contact attempts per case or per complete
  - Completes minus goal
Performance Dashboards

- Clear, concise presentation of KPIs, other important metrics
  - Just the essentials, in the best way for the user to understand quickly

- Graphical interfaces to production systems: balance of standardization, flexibility
  - Customization of displays in real time
  - Support for actions
  - Drill-down capability

- Drive decision-making, and the power that comes with access to a number of large databases, down to managers
Encoding Data for Rapid Perception

How many 5's are there?

192793774596113394741848211766685146
934813766623772889789992481442556688
178734549894544522789238165341929987
518225955234674128639626239174389497
The Power of Pre-attentive Processing

Now how many 5’s do you see?

192793774596113394741848211766685146
934813766623772889789992481442556688
178734549894544522789238165341929987
518225955234674128639626239174389497
Increasing Interest in Survey Dashboards

- Need to pull data from multiple sources
  - Paradata explosion
  - Decreasing response rates, increasing cost and quality pressure
  - Multiple modes
  - Responsive/Adaptive design

- Dashboards offer a solution
  - IT advances, increasingly rapid flow of information
  - Increased communication speed and modes
  - Survey professionals/managers’ skill requirements
Paradata Management

- Many potential data sources: interview or assessment timings, case status, record of calls, payroll and expense data, interview notes, interviewer characteristics, audio files, keystroke files, location data
- Some are very large (GPS data are Big Data)
- Some may be unstructured (audio files, interviewer notes)
- Some have complex structures (call record data – many records for one case, case may spawn other cases, and case status data are hierarchical, draw from questionnaire status)
- Dashboard must be a single source of truth
Evolution of Performance Dashboards at Westat
Origins at Westat

- My perspective: face-to-face household surveys
- Recognizing paradata challenges (2005-2010)
  - Separate data bases, data flows for production, cost, and quality, complex structure
- Developing a solution for paradata structural complexity ("the Cube") (2011-2015)
  - Reports for field supervisors
- Clinical Trials Support Unit (CTSU) dashboard requirement (2014)
2015 Development Schedule

- **January/February**
  - Developing common language
  - Agreeing on general approach
  - Defining the user
  - Identifying metrics

- **March/April: Parallel tracks**
  - Standing up the paradata
  - Creating views into the data, using M3 and dot.net

- **May/June: testing, iterating**
Pd3 Metrics

- Interviewer hours per completed interview (HPC)
- Response rates (RR) by sample type
- Completes compared to goals
- By interviewer: Overall quality score for first interview coded compared to next interview coded after feedback
- Interviewer work status, location
- For alerts: Interviews completed at unusual times, or too short, or without consent to audio-record (signaling potential for falsification)
Field Supervisor Dashboard Layout, March ‘15
Initial Deployment, July ‘15
Post-Deployment

- July ‘15: Trained about a dozen field supervisors
- October ‘15: Debriefed supervisors, began dissemination to other projects, and development of v2
- May ‘16: Christened “Paradata Discovery Dashboard” (Pd3)
- Branched out in past year to develop web and telephone versions, client versions, short course, get experience into the literature
Rapid Feedback

- Dashboard speeds information flow
- Westat research program has found powerful effects on data quality when interviewers get verbal and written feedback within 72 hours of interview
- Enhanced sense of belonging to a team dedicated to quality improvement
- Virtuous cycle
- Can also act as a deterrent
Managing Quality alongside Production, Cost

Dashboards

- Push responsibility down to the manager for making tradeoffs that include specific quality elements
- Can lead to insights about tradeoffs because data quality metrics are displayed alongside production and cost metrics
- Can highlight various dimensions of quality, and give them more prominence for the manager
- Can be an important tool for reducing total survey error
Case Study:
US PIAAC Dashboards
PIAAC
A Multi-Cycle International Programme

- Examines a range of basic skills in the information age
- Assesses these adult skills consistently across participating countries

- The first cycle of PIAAC
  - 24 countries participated in 2011–12 (Round 1)
  - 9 countries participated in 2014–15 (Round 2)
  - 5 countries are participating in 2017–18 (Round 3)
US PIAAC

- Participation in all three rounds of the first cycle
  - Round 1 sample size ≈ 5,000
  - Round 2 sample size ≈ 3,600
  - Round 3 sample size ≈ 3,800
Why Real-Time Monitoring of Data Collection Process Matters

Sample of addresses

Interview Paths

Correct address

Correct SP

BQ

BQ bias

Assessment

Assessment bias

Refusal

Wrong address

Wrong SP

Refusal

Refusal

Sample data

Contaminated data - Interviewer influence

No data – Falsified data
Why Real-Time Monitoring of Data Collection Process Matters

Interview Paths

Correct address

Wrong address

Correct SP

Wrong SP

BQ

BQ bias

Assessment bias

Assessment

Refusal

Sample data

Contaminated data - Interviewer influence

No data – Falsified data
Why Real-Time Monitoring of Data Collection Process Matters

Interview Paths

Correct address

Correct SP

Wrong address

Wrong SP

Refusal

Refusal

BQ bias

Assessment bias

Assessment

Assessment

Sample data

Contaminated data - Interviewer influence

No data – Falsified data
Why Real-Time Monitoring of Data Collection Process Matters

Correct address → Correct SP → BQ bias → Refusal

Correct address → Correct SP → Assessment bias → Refusal

Wrong address → Wrong SP → Refusal

Interview Paths

Sample data
Contaminated data - Interviewer influence
No data – Falsified data
Data Collection QC
US PIAAC Round 1 and Round 2

- Web-based interactive Case Management System (CMS) to
  - Manage case information
  - Record disposition codes
  - Review interviewer productivity
  - Monitor overall production

- Data collection monitored through manual inspection of a large number of reports
Data Collection QC
US PIAAC Round 1 and Round 2 (2)

- Reports followed PIAAC Standards and Guidelines on falsification detection and other QC
  - Each instrument duration
  - Time between interviews
  - Interviews conducted very late/very early
  - Number of interviews per day
  - Monitoring data quality
  - Interviewer productivity (highest producing interviewers)
  - Validation
  - Observations/audio recording
  - …
US PIAAC Round 1&2 QC Monitoring
Data Collection QC
US PIAAC Round 3

Switched to managing and monitoring the progress of data collection in the field using Westat’s new system

- Mobile phones
- GPS tracking system
- Dashboard

Exception

- CARI (Computer Audio-Recorded Interviewing) not used because the VM does not have the capability to capture voice data
- Timing only monitored at the instrument level
  - Item timing is not accessible during data collection
Application of Mobile Phones in US PIAAC

- iPhone used to increase efficiency
  - Record field work and travel time
  - Enter record of contacts
  - Allow GPS tracking
US PIAAC Dashboards

- Regional and home office manager dashboard
  - Seven portlets
  - Interviewer window

- Field supervisor dashboard
  - Seven portlets
  - Only showing the supervisor’s region
    - Except productivity portlet shown for all regions
  - Interviewer window
PIAAC Dashboard Portlets - Example

Hours Per Complete By Region

<table>
<thead>
<tr>
<th>Region</th>
<th>Hours Per Complete</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region 1</td>
<td>122</td>
</tr>
<tr>
<td>Region 2</td>
<td>170</td>
</tr>
<tr>
<td>Region 3</td>
<td>215</td>
</tr>
<tr>
<td>Region 4</td>
<td>171</td>
</tr>
<tr>
<td>Region 5</td>
<td>178</td>
</tr>
<tr>
<td>Region 6</td>
<td>205</td>
</tr>
</tbody>
</table>
PIAAC Dashboard Portlets – Example Hours Per Complete For Region 1
Data Collection Monitoring in Round 3 – Results

- Significant time and cost savings for field supervisors, regional directors and home office management to
  - Review status
  - Review productivity
  - Identify falsifiers

- Automated alerts
  - Enables rapid response to crisis in the field (reduces burden and costs)
  - Enables real-time monitoring of falsifiers
    - Reduces the burden of re-fielding falsified cases
Future Directions
2\textsuperscript{nd} Cycle of PIAAC and Beyond:

*Establishing an automated process that further minimizes interviewer error and falsification*

- US PIAAC experience shows a significant improvement in data quality, at reduced monitoring costs, using:
  - Mobile phones
  - GPS tracking system
  - Data collection dashboard
2\textsuperscript{nd} Cycle of PIAAC and Beyond:

\textit{Establishing an automated process that further minimizes interviewer error and falsification (2)}

- Other Westat experiences show CARI to be a critical source for improving data quality and validation
- Monitoring item-completion time also an important tool for identifying interviewer effects and falsifiers
CARI and Time Data – Critical Monitoring Tools for QC of Assessments

- **CARI**
  - Unobtrusive (unlike tape recording)
  - Applies to 100% of cases
  - Great value for observing interviewing flaws
  - Great value for detecting falsification
  - Tag recording to match specific items in an instrument

- **Time data**
  - Additional portlets can be created to show item-completion time data patterns and outliers, using statistical regression models
  - Alert portlets can send alerts to supervisors in a real-time basis
2\textsuperscript{nd} Cycle of PIAAC and Beyond: Data Monitoring Process – Data Requirements

- An automated process requires
  - \textit{Case Management System}

- An ideal automated process requires real-time access to
  - \textit{Voice Data - CARI}
  - \textit{Time data}

- An ideal system will include
  - \textit{GPS tracking system – mobile app}
  - \textit{Dashboard}
Catching Data Collection Errors In Real-Time

Sample of addresses

Correct address

Wrong address

Correct SP

Wrong SP

BQ bias

Assessment

Refusal

Alert portlet

Time data

GPS

CARI

Interview Paths

Production portlet

Sample Data

- Contaminated data - interviewer influence

- No data – Falsification

Dashboard portlets
Summary and Conclusions
Summary and Conclusions

- Performance dashboards
  - A visual display of the most important information on a single screen
  - Provides a real-time monitoring of the progress of data collection and signals unusual outcomes
- Evolution of performance dashboards at Westat
- Application of dashboards during US PIAAC Round 3 data collection
  - Significant monitoring time and cost savings
- Sample monitoring in Future Cycles – A Wish List
  - Case Management System
  - Voice and time data
  - GPS tracking and dashboards
Thank You

leylamohadjer@westat.com