Foundations of
NAEP Process Data

COPAFS Quarterly Meeting
Dec 4, 2020
Emmanuel Sikali, Ph.D.
Senior Research Scientist/Mathematical Statistician, NCES
Agenda

• Overview NAEP and NAEP Assessments
• Origins of the NAEP Process Data
• Process Data Research Examples
  • Students’ Test Taking Behavior
  • Item/Test Development and Scoring
  • Assessment Features
  • Assessment Accommodations
• Process Data Community
Overview: NAEP and NAEP Assessments
National Assessment of Educational Progress (NAEP)

- NAEP is a congressionally mandated assessment and serves as an integral part of our nation’s evaluation of the condition and progress of education.
- NAEP is the largest nationally representative and continuing assessment of what America's students know and can do in various subject areas.
- NAEP is required under the Elementary and Secondary Education Act of 1965, which was reauthorized as the Every Student Succeeds Act of 2015.
- The first national assessments were held in citizenship, science, and writing to 9-, 13-, and 17-year-olds in 1969.
Subjects Assessed for NAEP

- NCES administers NAEP assessments **in public and nonpublic (private)** schools across the nation.

- **Four subjects**— mathematics, reading, science, and writing at grades 4, 8, and 12—are assessed most frequently and reported at the national level (mathematics and reading grades 4 and 8 reported at the state and large urban district level, as well).
Sampling for NAEP

- NAEP is designed to report results at the national and state level, as well as for selected urban districts by creating a sampling frame.
## WE ARE LIVING IN A DIFFERENT ERA

<table>
<thead>
<tr>
<th></th>
<th>Past</th>
<th>Present</th>
<th>Future</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Item development</strong></td>
<td>Labor intensive</td>
<td>Labor intensive</td>
<td>Automatized</td>
</tr>
<tr>
<td><strong>Item types</strong></td>
<td>Generic</td>
<td>Enhanced</td>
<td>Real-life</td>
</tr>
<tr>
<td><strong>Test design</strong></td>
<td>Static</td>
<td>Semi-static</td>
<td>Data-driven</td>
</tr>
<tr>
<td><strong>Test assembly</strong></td>
<td>Labor intensive</td>
<td>Semi-automatized</td>
<td>Automatized</td>
</tr>
<tr>
<td><strong>Accessibility</strong></td>
<td>Limited</td>
<td>Universal design</td>
<td>Adaptive</td>
</tr>
<tr>
<td><strong>Timing</strong></td>
<td>Not measurable</td>
<td>Measured</td>
<td>Used</td>
</tr>
<tr>
<td><strong>Pathways</strong></td>
<td>Not observable</td>
<td>Observable</td>
<td>Modeled</td>
</tr>
<tr>
<td><strong>Validity</strong></td>
<td>Content/Corr based</td>
<td>Construct based</td>
<td>Process based</td>
</tr>
<tr>
<td><strong>Feedback</strong></td>
<td>Summative</td>
<td>Summative</td>
<td>Diagnostic</td>
</tr>
</tbody>
</table>
Origins of the NAEP Process Data
NAEP Process Data Snapshot

NAEP digitally based assessment

Students’ interactions are logged

Rich data source is formatted for further purposes

Examinees’ real time interactions with the digital test system environment that are recorded in the background as timestamped events.
Design of NAEP Mathematics Assessments

- Tutorial: 6-9 min
- Two cognitive blocks (Only mathematics items): 30 min (90 min) + 30 min (90 min)
- Student survey questionnaire block (Demographic and mathematics-specific survey questions): 15 min

Process Data
Which of the following three-dimensional shapes is a sphere?

A  

B  

C  

D  

What is the volume, in cubic centimeters, of a sphere with a radius of 5 centimeters? Express your answer in terms of \( \pi \).
Benefits and Challenges of NAEP Process Data

**OPPORTUNITIES**

- **Insights into**
  - students’ testing experiences,
  - problem solving behaviors (e.g., guessing, skipping pattern of items)
  - misconceptions
  - metacognitive processes (e.g., item revisits)
  - motivation, persistence, and engagement

- **Data supported operational decisions**
  - item development, analysis and selection
  - questionnaire development and validation
  - block and test assembly optimization
  - understanding test administration conditions

- Insights into **learners’ needs, accommodation use and effect**

- **Enhancing the communication** of what assessments measure

- **Modeling** cognitive and behavioral processes, advancing psychometric methods or building new IRT models, developing a framework for process data use

- **Advancing** research, item/test development, reporting, teaching/learning practices, and decision making

**CHALLENGES**

- Collecting data **with more expert input** may increase potential utility of data

- Process data are **noisy, complex**, and require detailed exploration

- **Standards** are under construction

- **Field has limited access** to NAEP process data which inhibits forming a research community centered around process data

- **Limited platforms/systems** that can be integrated across various assessments
Process Data Workflow

- Data
- Data Management
- Security & Privacy
- Analytics
- Viz
NAEP Process Data
### Data Infrastructure

```xml
<?xml version="1.0" encoding="utf-8"?>
<assessmentResult>

<context>
  <sessionIdentifier sourceId="Database Version 666" softwareVersion="6.0.28.4501" superVersion="3.5.1" chromeExtension="3.1.5">
    <bookletNumber></bookletNumber>
    <assignedForm>P213</assignedForm>
  </context>
  <testResult assessmentClear="2019" subjectName="Mathematics" assessedGroup="Grade 8" datestamp="2019-03-01">
    <outcomeVariable cardinality="record" interpretation="AdministrationCode">
      <value fieldIdentifier="AdministrationCode" baseType="integer" value="12" />
    </outcomeVariable>
    <outcomeVariable cardinality="record" interpretation="OriginalSession">
      <value fieldIdentifier="OriginalSessionDescription" baseType="string" originalsession="In session part time" />
    </outcomeVariable>
    <outcomeVariable cardinality="record" interpretation="Break">
      <value fieldIdentifier="Break" baseType="string" value="Break" />
    </outcomeVariable>
    <outcomeVariable cardinality="single" baseType="string" interpretation="TeacherNumber">
      <value value="06" />
    </outcomeVariable>
  </testResult>

  <itemResult accessionNumber="Adjust" itemType="Adjustment" blockCode="1919MA6AXX0X0X03EX">
    <outcomeVariable cardinality="single" interpretation="Enter Item">
      <value fieldIdentifier="EventTime" baseType="DateTime" value="2019-02-26T16:25:11" />
    </outcomeVariable>
  </itemResult>

  <itemResult accessionNumber="Intro-M8" itemType="Tutorial" blockCode="1919MA2A01TXX00X03EX">
    <outcomeVariable cardinality="single" interpretation="Enter Item">
      <value fieldIdentifier="EventTime" baseType="DateTime" value="2019-02-26T16:25:11" />
    </outcomeVariable>
  </itemResult>

  <itemResult accessionNumber="Intro-M8" itemType="Tutorial" blockCode="1919MA2A01TXX00X03EX">
    <outcomeVariable cardinality="single" interpretation="Enter Item">
      <value fieldIdentifier="EventTime" baseType="DateTime" value="2019-02-26T16:25:11" />
    </outcomeVariable>
  </itemResult>

  <itemResult accessionNumber="Intro-M8" itemType="Tutorial" blockCode="1919MA2A01TXX00X03EX">
    <outcomeVariable cardinality="single" interpretation="Enter Item">
      <value fieldIdentifier="EventTime" baseType="DateTime" value="2019-02-26T16:25:11" />
    </outcomeVariable>
  </itemResult>

  <itemResult accessionNumber="Intro-M8" itemType="Tutorial" blockCode="1919MA2A01TXX00X03EX">
    <outcomeVariable cardinality="single" interpretation="Enter Item">
      <value fieldIdentifier="EventTime" baseType="DateTime" value="2019-02-26T16:25:11" />
    </outcomeVariable>
  </itemResult>
</assessmentResult>
```

<table>
<thead>
<tr>
<th>studentID</th>
<th>originalOrder</th>
<th>accessionNumber</th>
<th>itemType</th>
<th>blockCode</th>
<th>interpretation</th>
<th>time Stamp</th>
<th>extendedInfo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Adjust</td>
<td>Adjustment</td>
<td>1919MA6AXX0X0X03EX</td>
<td>Enter Item</td>
<td>2019-02-19T16:25:11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Adjust</td>
<td>Adjustment</td>
<td>1919MA6AXX0X0X03EX</td>
<td>Change Theme</td>
<td>2019-02-19T16:25:11</td>
<td>blackBeige</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Adjust</td>
<td>Adjustment</td>
<td>1919MA6AXX0X0X03EX</td>
<td>Change Theme</td>
<td>2019-02-19T16:25:11</td>
<td>whiteBlack</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Adjust</td>
<td>Adjustment</td>
<td>1919MA6AXX0X0X03EX</td>
<td>Next</td>
<td>2019-02-19T16:25:11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Adjust</td>
<td>Adjustment</td>
<td>1919MA6AXX0X0X03EX</td>
<td>Exit Item</td>
<td>2019-02-19T16:25:11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Intro-M8</td>
<td>Tutorial</td>
<td>1919MA2A01TXX00X03EX</td>
<td>Enter Item</td>
<td>2019-02-19T16:25:11</td>
<td></td>
<td>AudioStarted-ToolInt</td>
</tr>
<tr>
<td>7</td>
<td>Intro-M8</td>
<td>Tutorial</td>
<td>1919MA2A01TXX00X03EX</td>
<td>Media Interaction</td>
<td>2019-02-19T16:25:11</td>
<td>AudioComplete-ToolInt</td>
<td></td>
</tr>
</tbody>
</table>
## Example of Captured Events [Mathematics]

<table>
<thead>
<tr>
<th>Event</th>
<th>Event</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Back</td>
<td>Eliminate Choice</td>
<td>Move Calculator</td>
</tr>
<tr>
<td>Calculator Buffer</td>
<td>Equation Editor Button</td>
<td>Next</td>
</tr>
<tr>
<td>Change Theme</td>
<td>Erase</td>
<td>Open Calculator</td>
</tr>
<tr>
<td>Clear Answer</td>
<td>Exit Item</td>
<td>Open Equation Editor</td>
</tr>
<tr>
<td>Clear Choice</td>
<td>Hide Timer</td>
<td>Receive Focus</td>
</tr>
<tr>
<td>Click Progress Navigator</td>
<td>Highlight</td>
<td>Scratchwork Draw Mode On</td>
</tr>
<tr>
<td>Close Calculator</td>
<td>Horizontal Item Scroll</td>
<td>Scratchwork Erase Mode On</td>
</tr>
<tr>
<td>Close Equation Editor</td>
<td>Increase Zoom</td>
<td>Scratchwork Highlight Mode On</td>
</tr>
<tr>
<td>Decrease Zoom</td>
<td>Leave Section</td>
<td>TextToSpeech</td>
</tr>
<tr>
<td>Draw</td>
<td>Lose Focus</td>
<td>Vertical Item Scroll</td>
</tr>
<tr>
<td>DropChoice</td>
<td>Math Keypress</td>
<td></td>
</tr>
</tbody>
</table>
## Captured vs. Derived Variables

<table>
<thead>
<tr>
<th>Captured</th>
<th>Derived</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student Identifier</td>
<td>Cumulative time</td>
</tr>
<tr>
<td>Block Code</td>
<td>Number of visits (e.g., 1,2,3,4 )</td>
</tr>
<tr>
<td>Accession Number</td>
<td>Calculator Use (yes-no)</td>
</tr>
<tr>
<td>Item Type Code</td>
<td>Response change (e.g., A-&gt;B)</td>
</tr>
<tr>
<td>Observable Type</td>
<td>..</td>
</tr>
<tr>
<td>Extended Info</td>
<td></td>
</tr>
<tr>
<td>Timestamp</td>
<td></td>
</tr>
</tbody>
</table>
Beyond Data

Data Management
Data Security
Data Privacy
Data Quality

PROCESS DATA:
Data Request, Access, Storage, and Disposal Procedure

PROCESS DATA:
Data Quality Guideline
Gaining Insights: NAEP Process Data
Today’s Examples

Students’ Test Taking Behavior
Item/Test Development & Scoring
Assessment Features
Assessment Accommodations
Students’ Test Taking Behavior: Disengagement

Our Research:
Can finite mixture modelling techniques be used on response time to identify distinct groups of students with different testing behaviors, such as disengagement and speededness?

Key Findings:
Distinct behaviors can be identified. Meaningful and plausible interpretations can be made about the identified groups.

Item/Test Development & Scoring: Non-Response Rates

Our Research:
Explore how process data helps to evaluate the appropriateness of the conventional definitions for non-response and omission

– Do non-response and omission rates differ between scoring files (conventional) and process data?
– Can we identify a threshold between non-response and omit for each item using response time?

Key Findings:
Non-response rates between student scoring file and process data differ

<table>
<thead>
<tr>
<th>Item sequence</th>
<th>Item type</th>
<th>Response Time</th>
<th>Response File Coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MatchMS</td>
<td>69.44</td>
<td>Incorrect</td>
</tr>
<tr>
<td>2</td>
<td>MCSS</td>
<td>20.26</td>
<td>B</td>
</tr>
<tr>
<td>3</td>
<td>Composite</td>
<td>108.44</td>
<td>Correct</td>
</tr>
<tr>
<td>4</td>
<td>FillInBlank</td>
<td>212.27</td>
<td>Incorrect</td>
</tr>
<tr>
<td>5</td>
<td>MCSS</td>
<td>45.39</td>
<td>A</td>
</tr>
<tr>
<td>6</td>
<td>MatchMS</td>
<td>82.32</td>
<td>Incorrect</td>
</tr>
<tr>
<td>7</td>
<td>MCMS</td>
<td>32.20</td>
<td>Partial</td>
</tr>
<tr>
<td>8</td>
<td>CompositeCR</td>
<td>348.30</td>
<td>Omitted</td>
</tr>
<tr>
<td>9</td>
<td>ZonesMS</td>
<td>85.45</td>
<td>Correct</td>
</tr>
<tr>
<td>10</td>
<td>CompositeCR</td>
<td>449.58</td>
<td>Partial</td>
</tr>
<tr>
<td>11</td>
<td>MatchMS</td>
<td>339.15</td>
<td>Not reached</td>
</tr>
<tr>
<td>12</td>
<td>CompositeCR</td>
<td>NA</td>
<td>Not reached</td>
</tr>
<tr>
<td>13</td>
<td>MCMS</td>
<td>NA</td>
<td>Not reached</td>
</tr>
<tr>
<td>14</td>
<td>CompositeCR</td>
<td>NA</td>
<td>Not reached</td>
</tr>
</tbody>
</table>
Assessment Accommodations: Extended Time Accommodation

**Our Research:**
Exploring extended time accommodation (ETA), by analyzing the relationship between ETA use and performance of students with ETA

**Key Findings:**
Only one-third of ETA students (35%) used extra time
On average, ETA students who used extra time scored 2 points higher than those ETA students who did not use extra time

Assessment Features: Calculator Use Study

Our Research:

• What computations students do with the calculators?

• How often do students use computation strategies or commit common errors identified by experts using the calculator?

Key Findings:

• Percentage of students who exactly follow anticipated solution strategies or errors vary.

• Most common computations were not exactly matching with the anticipations. Most non-matching computations were one character different than anticipated. Unanticipated errors were also found.

Process Data Community Building: 2019 NCME and 2020 AERA
Process Data Community Building: Special Interest Groups (SIGs)

**AERA:** Process Data in Digitally Based Assessments

Process data research is an emerging topic that has sparked growing interest due to its novelty and yet unexplored potential.

**NCME:** Big Data in Educational Measurement

(Join: [https://form.jotform.com/ncme/SIGIMIESIGNUP](https://form.jotform.com/ncme/SIGIMIESIGNUP))

Leverage the availability of big data from a variety of sources to inform the study of education and educational measurement.
NCES Process Data File

• Data file produced from
  • Assessment year 2017
  • Grade 8
  • Subject mathematics

For more information please contact iesdata.security@ed.gov
For your questions

Contact information: emmanuel.sikali@ed.gov