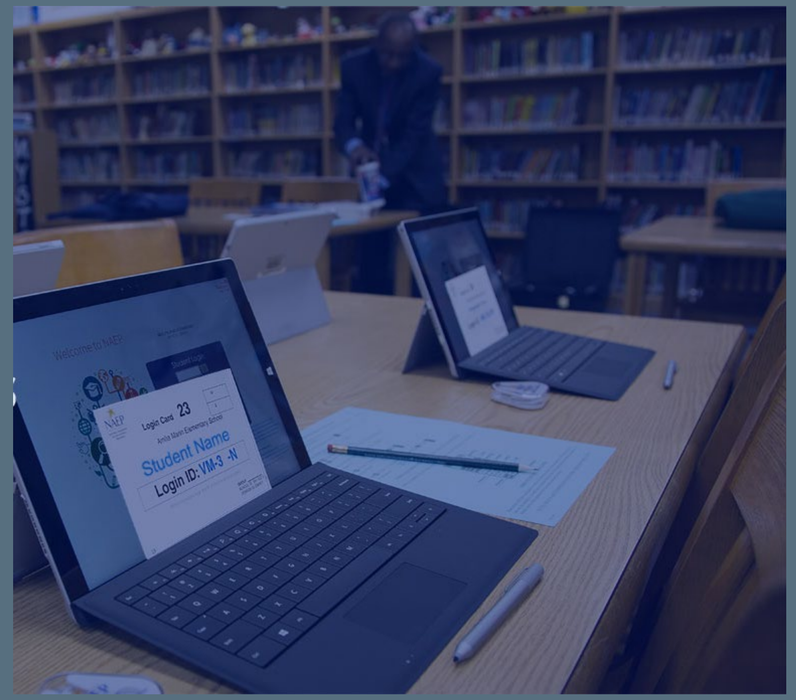
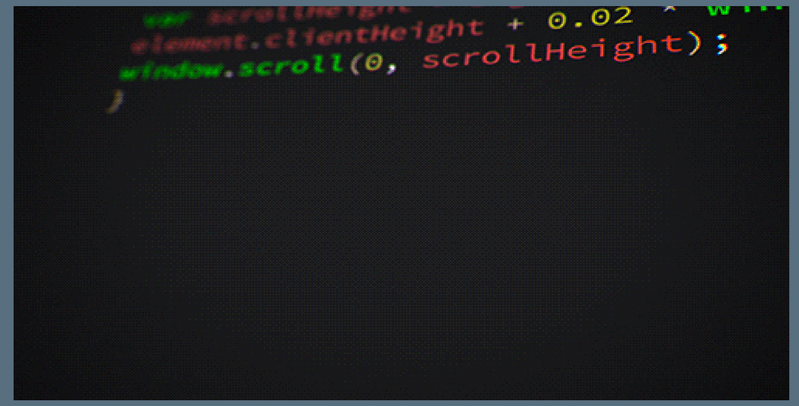
NAEP Process Data: Past, Present, and Future

Emmanuel Sikali, Ph.D.
Senior Research Scientist/Mathematical Statistician, NCES





IES Institute of Education Sciences

Agenda

- Overview
  - Introduction to NAEP assessments
  - Change of Assessment in Digital Era
- NAEP Process Data
  - Origin
  - Collection and Cleaning Process
- NAEP Process Data: Potential
  - 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)

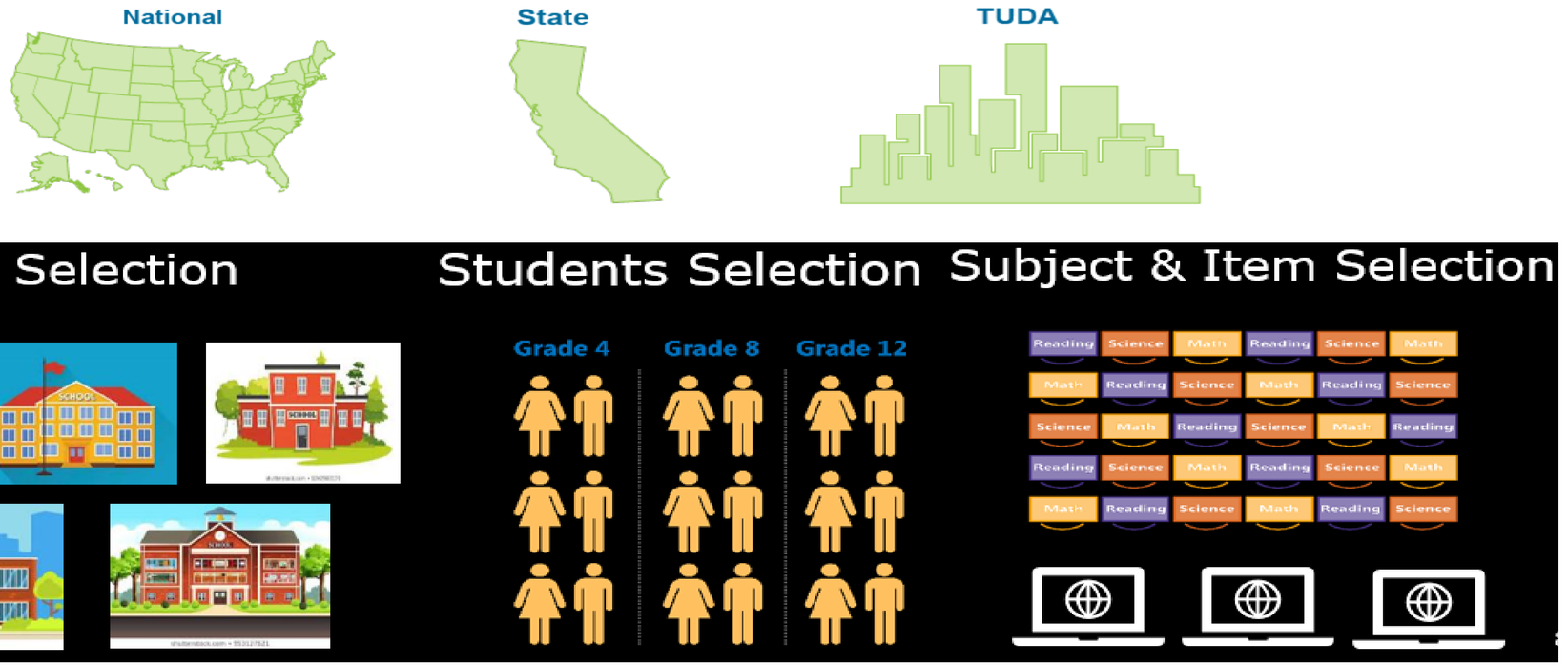
- NAEP is a **congressionally mandated** assessment and serves as an integral part of our nation's evaluation of the condition and progress of education.
- NAEP is the largest nationally representative and continuing assessment of **what America's students know and can do** in various subject areas.
- 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- are assessed most frequently and reported at the state and district level**.**

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.



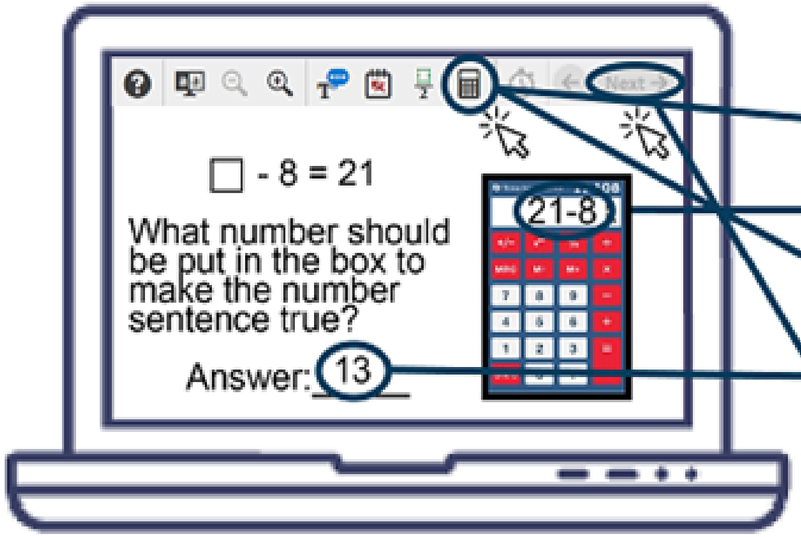
WE ARE LIVING IN A DIFFERENT ERA

	PAST	PRESENT	**FUTURE**
Item development	Labor intensive	Labor intensive	Automatized
Item types	Generic	Enhanced	Real-life
Test design	Static	Semi-static	Data-driven
Test assembly	Labor intensive	Semi-automatized	Automatized
Accessibility	Limited	Universal design	Adaptive
Timing	Not measurable	Measured	Used
Pathways	Not observable	Observable	Modeled
Validity	Content/Corr based	Construct based	Process based
Feedback	Summative	Summative	Diagnostic

Overview: NAEP Process Data

Process Data Overview

Digitally Based Assessment



Student Interactions Logged

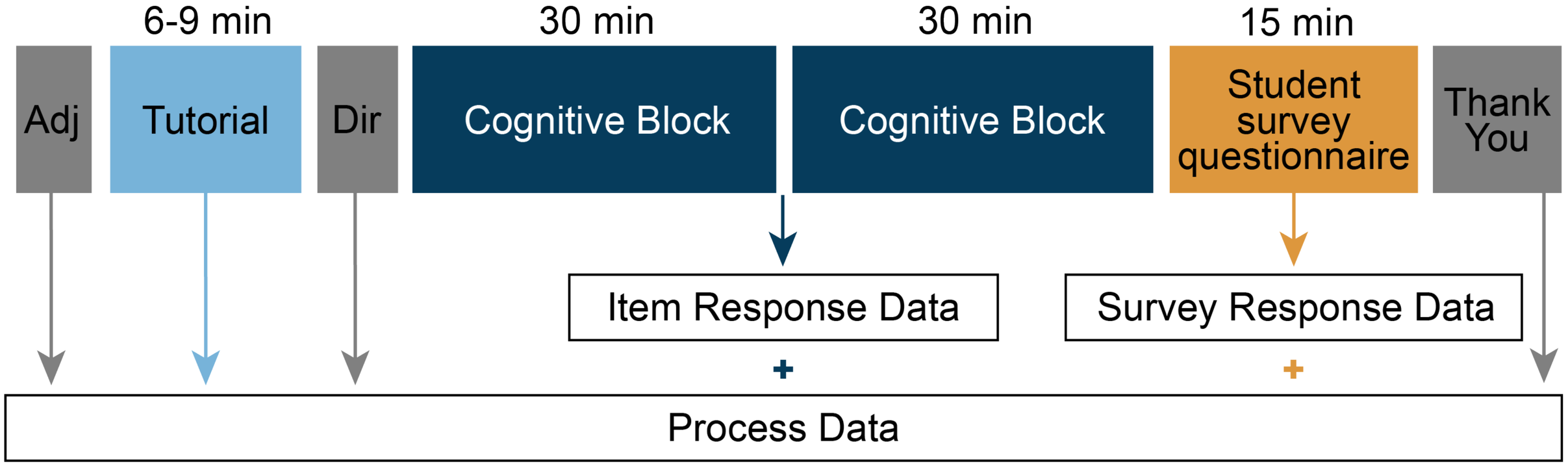
- ENTER ITEM
- OPEN CALCULATOR
- TYPE 21 - 8
- CLOSE CALCULATOR
- TYPE 13
- CLICK NEXT
- EXIT ITEM

Process Data Table

Block	Student ID	Item	Event	Time
MX	1001	1	Enter Item	1:01
MX	1001	1	Open Calculator	1:12
MX	1001	1	Type 21 - 8	1:19
MX	1001	1	Close Calculator	1:34
MX	1001	1	Type 13	1:19
MX	1001	1	Click Next	1:34
MX	1001	1	Exit Item	1:34
...	...	...	...	...

Examinees' real time interactions with the digital test system environment that are recorded in the background as timestamped events.

Data from NAEP Assessment



Progress

Next

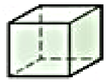
1 2 3 4 Review

System tools

Navigation Panel

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

A 

B 

C 

D 

Clear Answer

What is the volume, in cubic centimeters, of a sphere with a radius of 5 centimeters? Express your answer in terms of π.

[ ] cubic centimeters

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 is **noisy, complex**, and requires 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

NAEP Process Data: Origin

Logging Data

- Data structure: Structured data
- Collection of pre-decided actions
- Data Codebook : Not standardized

Checking and Cleaning

Data validation

- Correctness and consistency of data sets
  - Within data set
    - monotonically increasing time stamps
    - duplicated cases
  - Across data sets
    - consistency with item maps
- Security
  - item content, hate speech
- Merging with response data

Variable Derivation

Data Preparation for Analysis

- Derivation of variables
  - Time spent
  - Accommodation Use
  - UDE use
  - …

Data Infrastructure

```
<?xml version="1.0" encoding="utf-8"?>
<assessmentResult>
  <context>
    <sessionIdentifier sourceID="Database Version 686" softwareVersion="6.0.28.4501" superVersion="5.5.1" chromeExtension="3.1.5"
    <bookletNumber>[redacted]</bookletNumber>
    <assignedForm>M213</assignedForm>
  </context>
  <testResult assessmentYear="2019" subjectName="Mathematics" assessedGroup="Grade 8" datestamp="2019-03-01[redacted]
    <outcomeVariable cardinality="record" interpretation="AdministrationCode">
      <value fieldIdentifier="AdministrationCode" baseType="integer">12</value>
      <value fieldIdentifier="AdministrationCodeDescription" baseType="string">Original session - In session part time</value>
      <value fieldIdentifier="AdministrationNote" baseType="string">Break</value>
    </outcomeVariable>
    <outcomeVariable cardinality="single" baseType="string" interpretation="TeacherNumber">
      <value>06</value>
    </outcomeVariable>
  </testResult>
  <itemResult accessionNumber="Adjust" itemType="Adjustment" blockCode="1919MA6AXXAXXX03EX">
    <outcomeVariable cardinality="single" interpretation="Enter Item">
      <value fieldIdentifier="EventTime" baseType="dateTime">2019-02-28[redacted]value>
    </outcomeVariable>
  </itemResult>
```

studentID	originalOrder	accessionNumber	itemType	blockCode	interpretation	timeStamp	extendedInfo
[redacted]	1	Adjust	Adjustment	1919MA6AXXAXXX03EX	Enter Item	2019-02-19T[redacted]	
[redacted]	2	Adjust	Adjustment	1919MA6AXXAXXX03EX	Change Theme	2019-02-19T[redacted]	blackBeige
[redacted]	3	Adjust	Adjustment	1919MA6AXXAXXX03EX	Change Theme	2019-02-19T[redacted]	whiteBlack
[redacted]	4	Adjust	Adjustment	1919MA6AXXAXXX03EX	Next	2019-02-19T[redacted]	
[redacted]	5	Adjust	Adjustment	1919MA6AXXAXXX03EX	Exit Item	2019-02-19T[redacted]	
[redacted]	6	Intro-M8	Tutorial	1919MA2A01TXXC00EX	Enter Item	2019-02-19T[redacted]	
[redacted]	7	Intro-M8	Tutorial	1919MA2A01TXXC00EX	Media Interaction	2019-02-19T[redacted]	AudioStarted-ToolInt1
[redacted]	8	Intro-M8	Tutorial	1919MA2A01TXXC00EX	Media Interaction	2019-02-19[redacted]	AudioComplete-ToolInt1

Example of Captured Events [Mathematics]

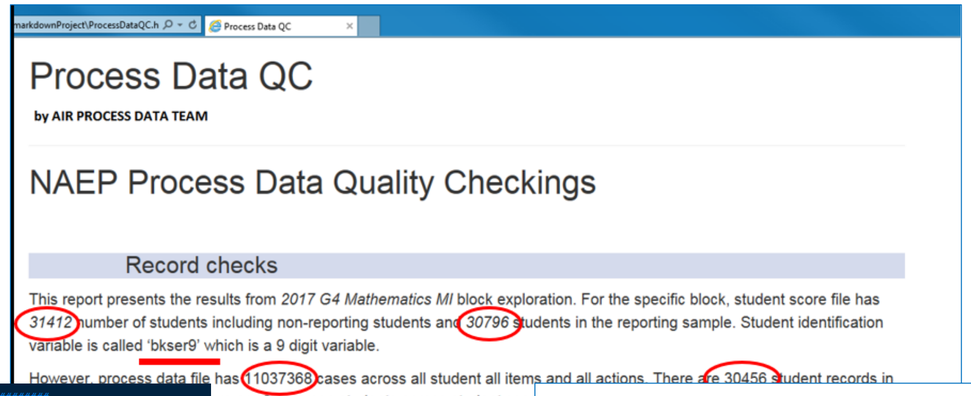
Event	Event	Event
Back	Eliminate Choice	Move Calculator
Calculator Buffer	Equation Editor Button	Next
Change Theme	Erase	Open Calculator
Clear Answer	Exit Item	Open Equation Editor
Clear Choice	Hide Timer	Receive Focus
Click Progress Navigator	Highlight	Scratchwork Draw Mode On
Close Calculator	Horizontal Item Scroll	Scratchwork Erase Mode On
Close Equation Editor	Increase Zoom	Scratchwork Highlight Mode On
Decrease Zoom	Leave Section	TextToSpeech
Draw	Lose Focus	Vertical Item Scroll
DropChoice	Math Keypress	

Captured vs. Derived Variables

Captured	Derived
Student Identifier	Cumulative time
Block Code	Number of visits (e.g., 1,2,3,4)
Accession Number	Calculator Use (yes-no)
Item Type Code	Response change (e.g., A->B)
Observable Type	….
Extended Info	
Timestamp	

Beyond Data

Data Management
Data Security
Data Privacy
Data Quality



```
##############################################################################
# READ process data
##############################################################################

# Get the names of the forms that include MC block
  # read booklet/block info
    bookmap<-data.table(read.csv(paste0(basedir,"/input/",
                                        "NAEP2017G8Bookmap.csv"),header=TRUE))

  # get rid of white spaces
    bookmap[,c("Block1","Block2"):=list(trimws(Block1),trimws(Block2))]
    bookmap[,FormNumber :=trimws(FormNumber)]
    #head(bookmap[,c(3,5,6,11,12)])

  # create block position
    bookmap[,BlockPosition:=ifelse(Block1==blocknames[1],"1",
                                   ifelse(Block2==blocknames[1],"2",NA))]
    # table(bookmap$Block1)

# Extract form names and block position information
  FormInfo_MC<-bookmap[Block1==blocknames[1] | Block2==blocknames[1],c("FormNumber","BlockPosition")]

# Read item/block data
  itemmap<-data.table(read.csv(paste0(paste0(basedir,"/input/",
                                             "NAEP2017G8Itemmap.csv")),header=TRUE))

# Get the list of files in the target directory
  FileNames<-list.files(path = inputdir_pro)
  FilesToRead<-FileNames[grep("^AM.*Rdata",FileNames)] # 54 forms

# Get MC data from each form by keeping only cognitive items data
  MCBlockData<-list()
```

PROCESS DATA:

Data Request, Access, Storage, and Disposal Procedure

PROCESS DATA:

Data Quality Guideline

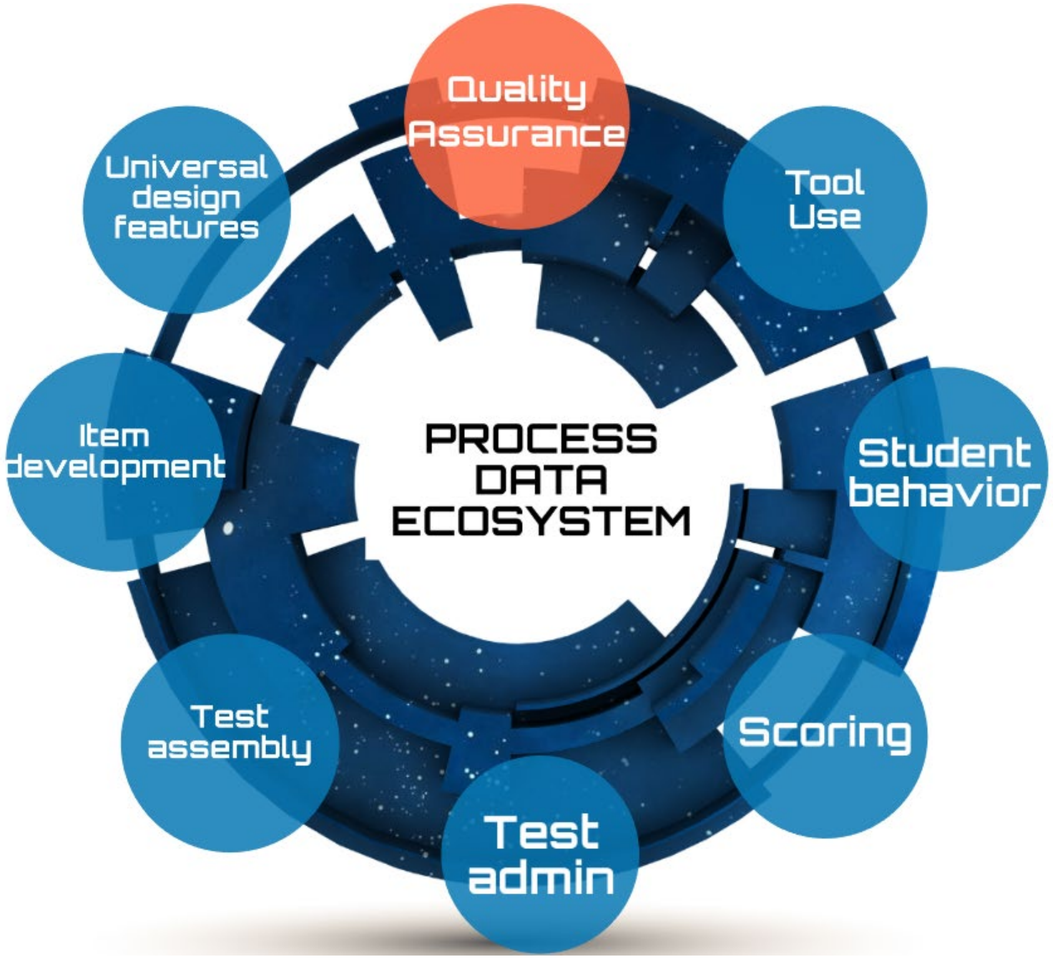
```
##############################################################################
# RT calculation
##############################################################################

# Calculate response time

  ResponseTime_MC <- data.frame(studentIDnew =character(), accessionNumber=character(),
                                RT=numeric(),stringsAsFactors = FALSE)

# Prepare timeStamp variable for calculation

  EntExtData_MC$timeStamp1<-gsub("T"," ", EntExtData_MC$timeStamp)
  EntExtData_MC$timeStamp2<-gsub("Z","", EntExtData_MC$timeStamp1)
  EntExtData_MC$timeStampnew<-as.POSIXct(EntExtData_MC$timeStamp2,format = "%Y-%m-%d  %H:%M:%OS")

# Loop over students and items:
```

NAEP Process Data: Potential

Data & Semantics

Identify & operationalize the variables
Generate defensible variables

Guide data collection
Provide base for the use of the data



Today's Examples

Students' Test Taking Behavior
Item/Test Development & Scoring
Assessment Features
Assessment Accommodations

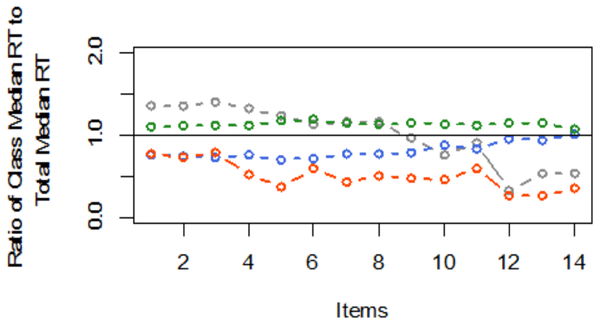
Students' Test Taking Behavior: Disengagement

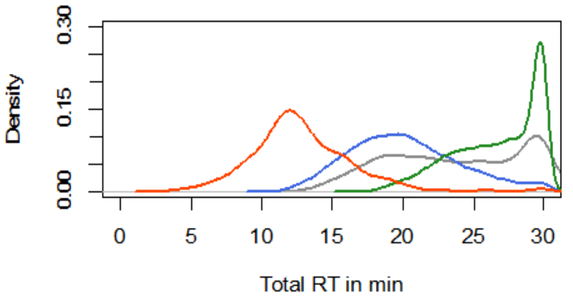
Our Research:

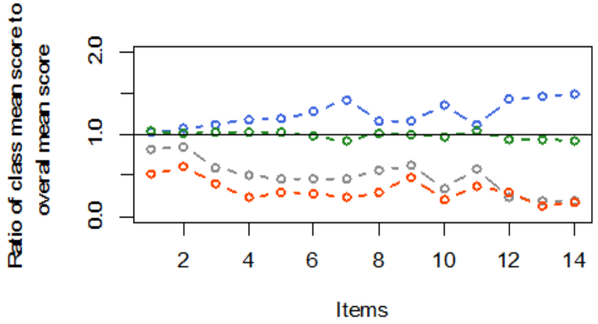
Can finite mixture modelling techniques be used on RT to identify distinct groups of examinees with different testing behaviors, including disengagement and speededness?

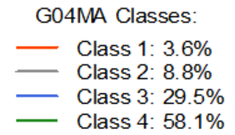
Key Findings:

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 NR and Omit

- Do NR and Omit rates differ between scoring files (conventional) and process data?
- Can we identify a threshold between NR and Omit for each item using RT?

Key Findings:

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

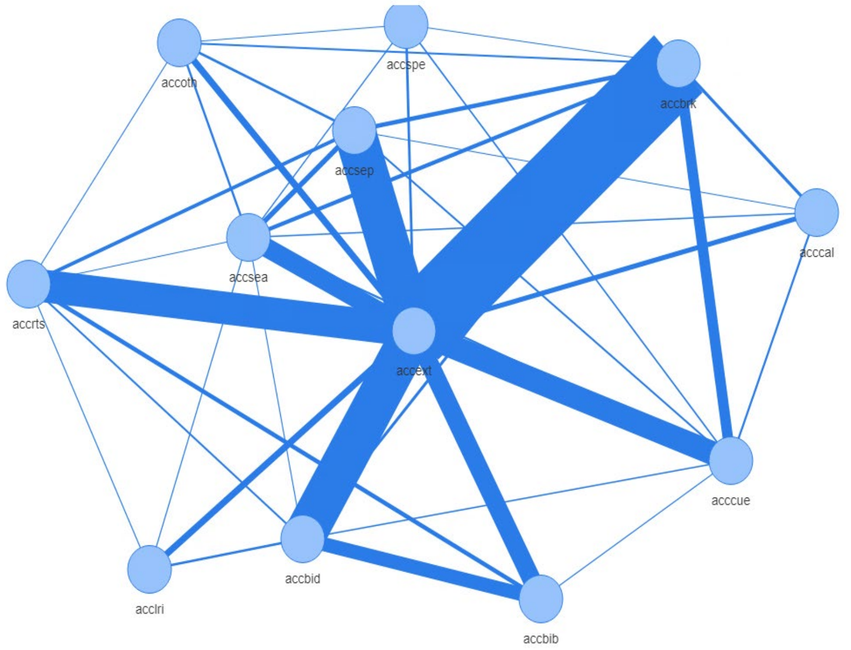
Item sequence	Item type	Response Time	Response File Coding
1	MatchMS	69.44	Incorrect
2	MCSS	20.26	B
3	Composite	108.44	Correct
4	FillInBlank	212.27	Incorrect
5	MCSS	45.39	A
6	MatchMS	82.32	Incorrect
7	MCMS	32.20	Partial
8	CompositeCR	348.30	Omitted
9	ZonesMS	85.45	Correct
10	CompositeCR	449.58	Partial
11	MatchMS	339.15	Not reached
12	CompositeCR	NA	Not reached
13	MCMS	NA	Not reached
14	CompositeCR	NA	Not reached

Digital Assessment and Accessibility

Our Research:
NAEP DBAs implement technologies to improve accessibility for all participants with different learning backgrounds.

Accessibility Feature	% Students allowed	% Students used	
		Accommodated	Non-accommodated
Extended Time	83*	32	NA
Text-to-Speech	All	45	30
Highlight	All	20	19
Draw	All	37	38
Eliminator Choice	All	33	37

Key Findings:
Students with accommodations use more of drawing and eliminator choice tools



Hicks, J., Circi, R., & Burhan, O. (2021).

Assessment Accommodations: Extended Time Accommodation

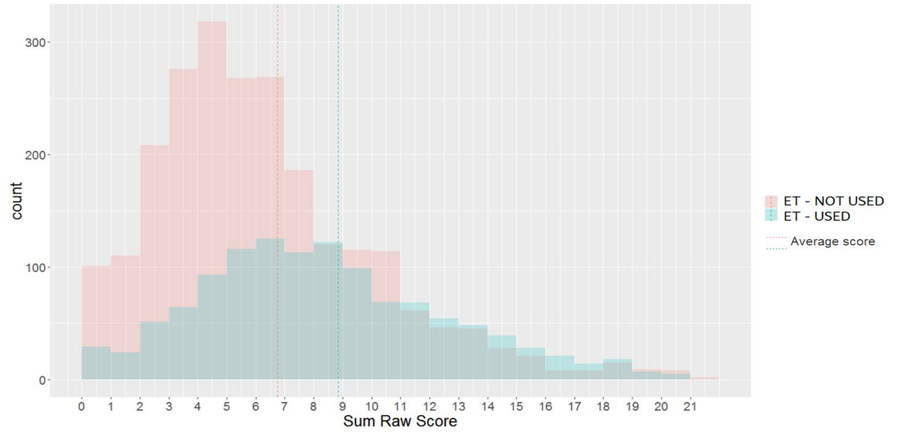
Our Research:

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 ET

On average, ETA students who used extra time scored 2 points higher than those ETA students who did not use extra time



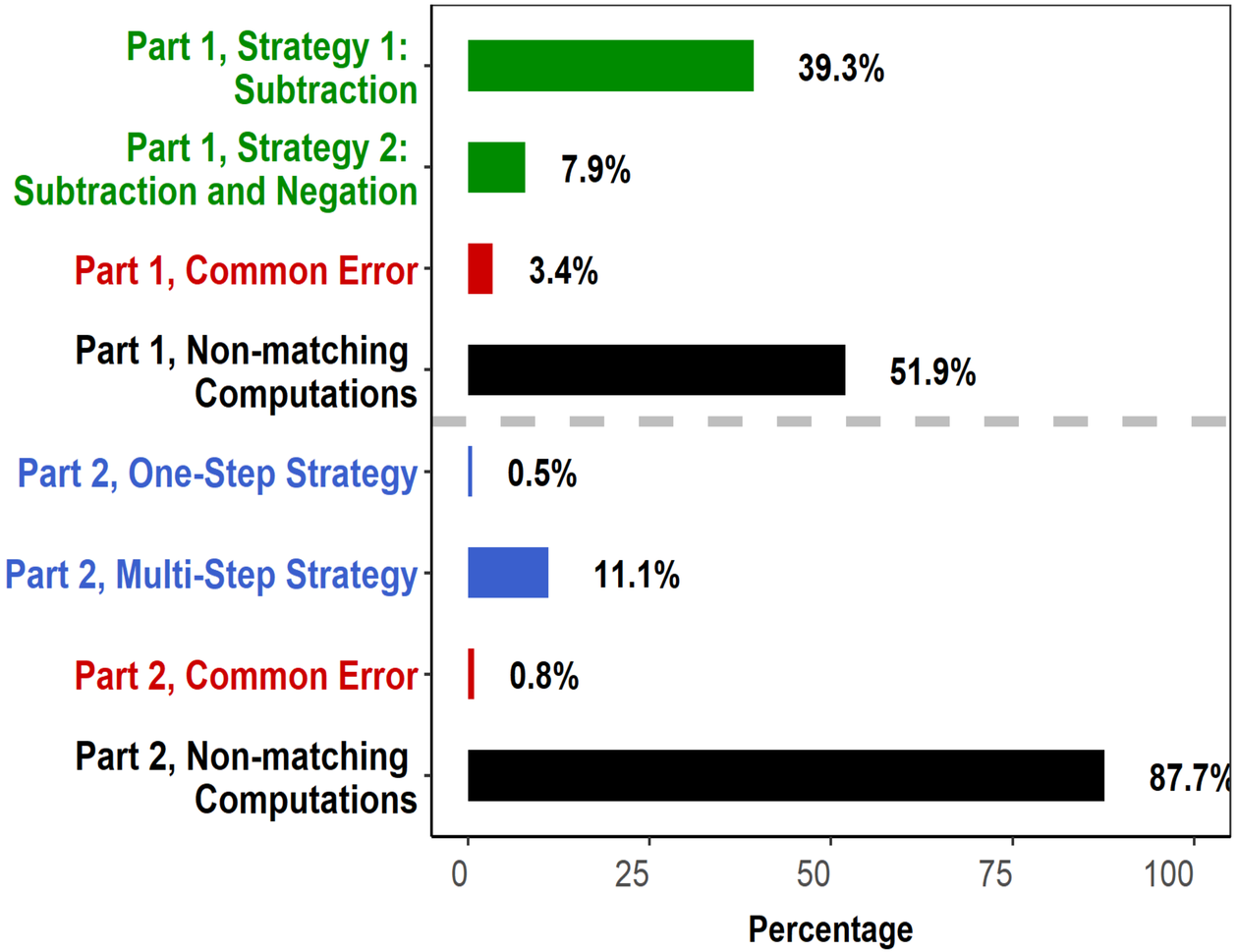
Assessment Features: Calculator Use Study

Our Research:

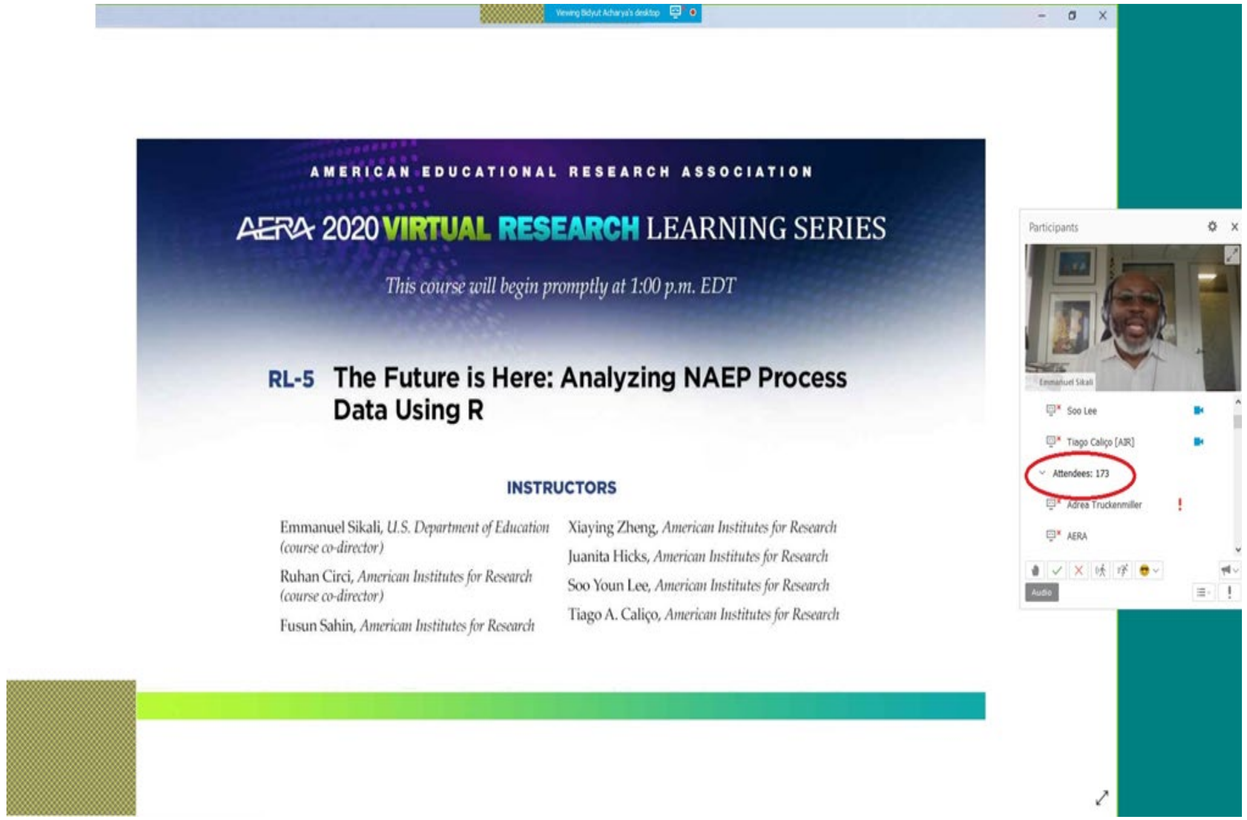
- What computations students do with the calculators?
- How often do students use computation strategies or commit common errors identified by experts using the calculator?

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 - 2021

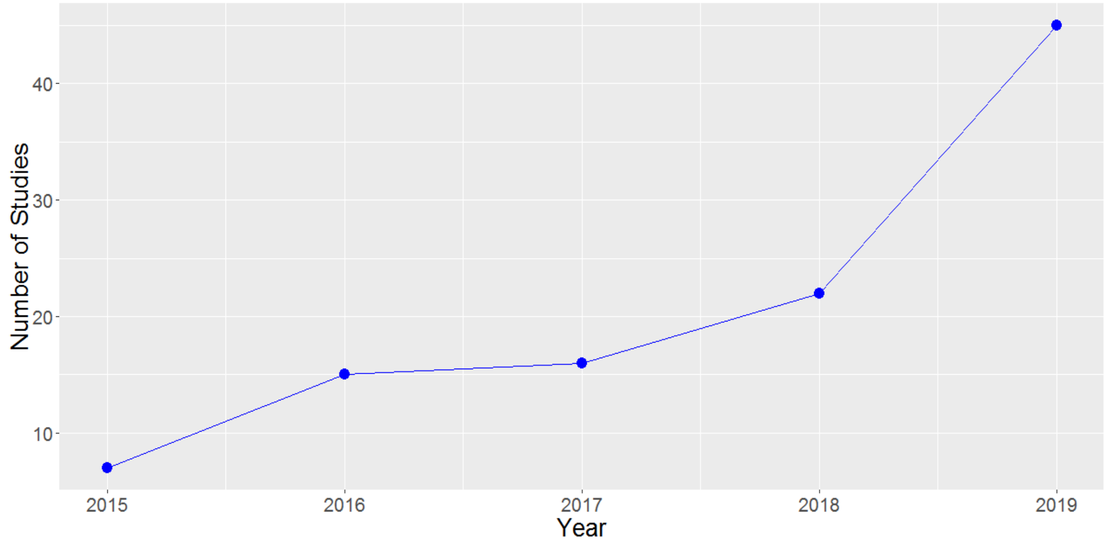


Process Data Community Building: Special Interest Groups (SIGs)

NCME: Big Data in Educational Measurement

(Join: 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



For your questions

Contact information:
emmanuel.sikali@ed.gov