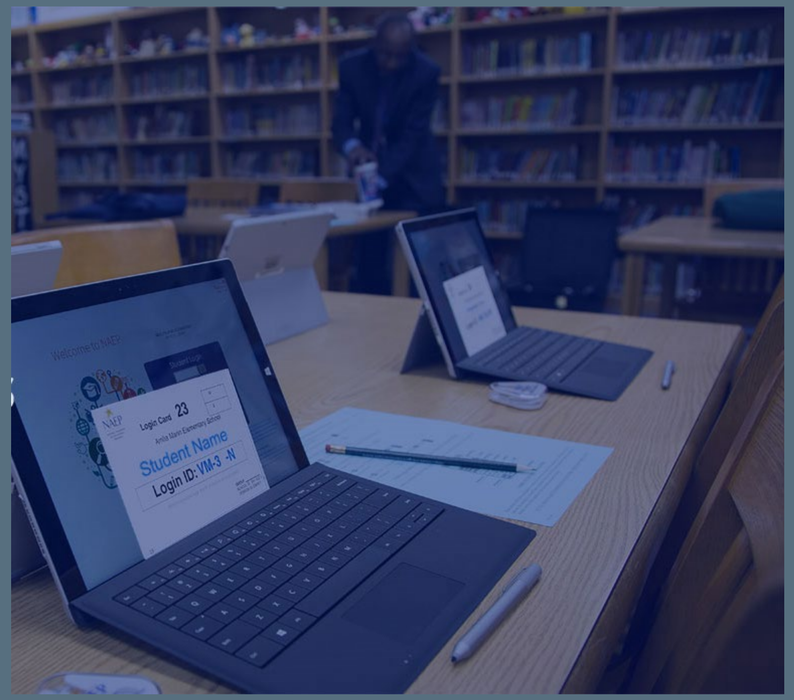
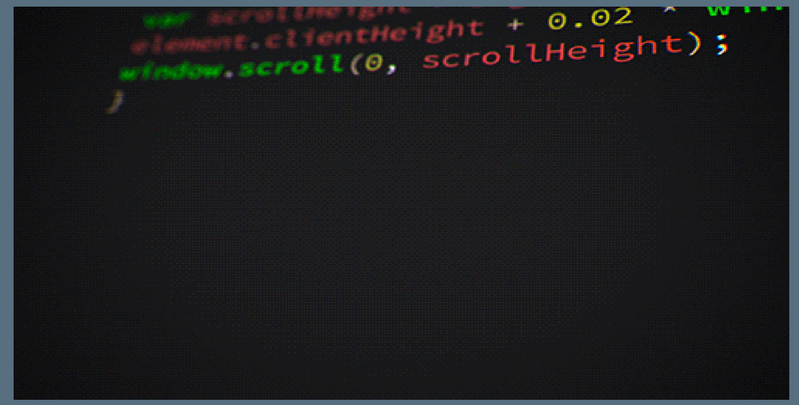
Foundations of NAEP Process Data





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

IES Institute of Education Sciences

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.

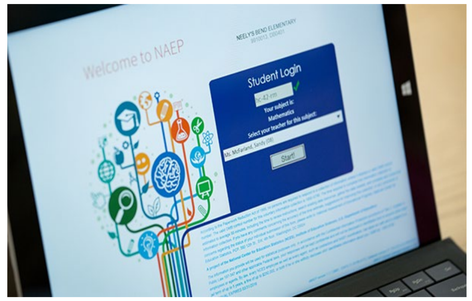
WE ARE LIVING IN A DIFFERENT ERA

	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

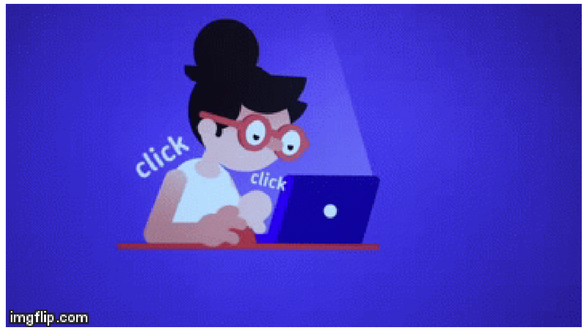
Origins of the NAEP Process Data

NAEP Process Data Snapshot

NAEP digitally based assessment



Students' interactions are logged

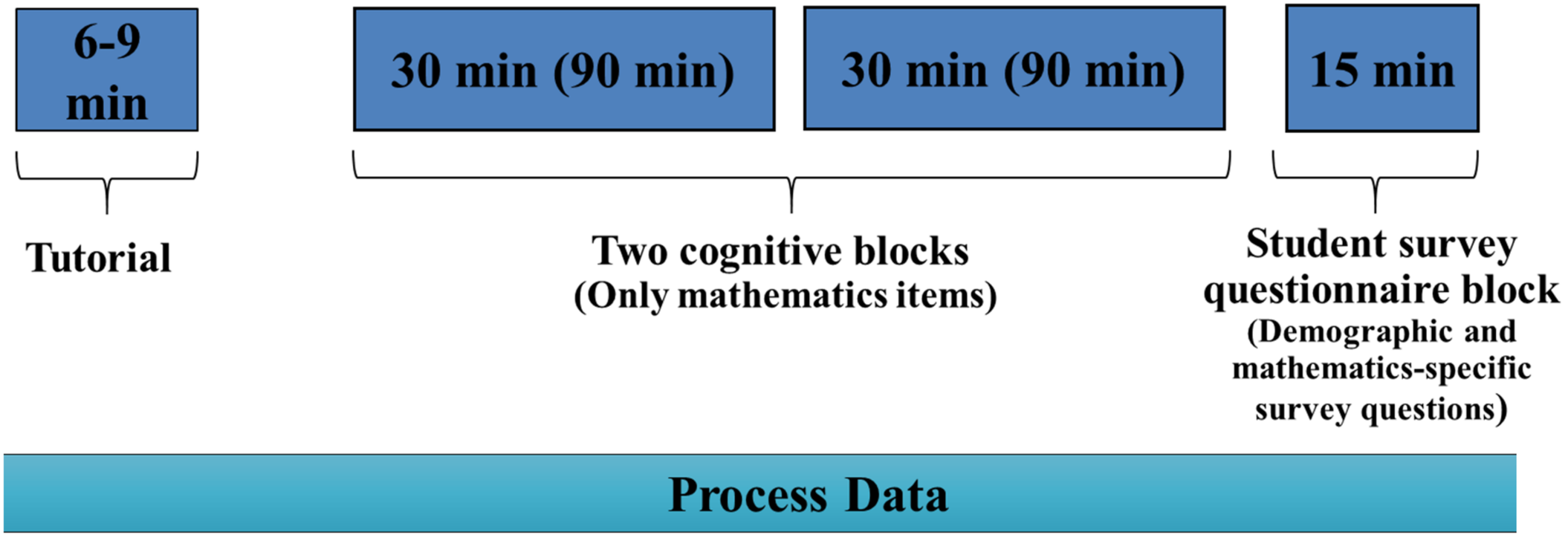


Rich data source is formatted for further purposes

PseudoId	BlockCode	AccessionNumber	ItemTypeCode	ObservableType	ExtendedInfo	Timestamp	Cumulative.Time
2209265315	1717MA2N03CLID30EX	VH356842	Directions	Enter Item	NULL	2017-03-06 10:00:00	0
2209265315	1717MA2N03CLID30EX	VH356842	Directions	Next	NULL	2017-03-06 10:00:10	10
2209265315	1717MA2N03CLID30EX	VH356842	Directions	Exit Item	NULL	2017-03-06 10:00:10	10
2209265315	1717MA2N03CLID30EX	VH266695	MCSS	Enter Item	NULL	2017-03-06 10:00:10	10
2209265315	1717MA2N03CLID30EX	VH266695	MCSS	Increase Zoom	125	2017-03-06 10:00:21	21
---	---	---	---	---	---	---	---
2209265315	1717MA2N03CLID30EX	SecTimeOut	TimeOutMessage	OK	null, ENG	2017-03-06 10:30:37	1837
2209265315	1717MA2N03CLID30EX	SecTimeOut	TimeOutMessage	Exit Item	NULL	2017-03-06 10:30:37	1837
2209265315	1717MA2N03CLID30EX	VH304553	MatchMS	Enter Item	NULL	2017-03-06 10:30:37	1837
2209265315	1717MA2N03CLID30EX	VH304553	MatchMS	Exit Item	NULL	2017-03-06 10:30:37	1837
2209265315	1717MA2N03CLID30EX	VH304553	MatchMS	Vertical Item Scroll	0, 0	2017-03-06 10:30:38	1838

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

Design of NAEP Mathematics Assessments



Progress

Next

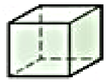
1 2 3 4 Review

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

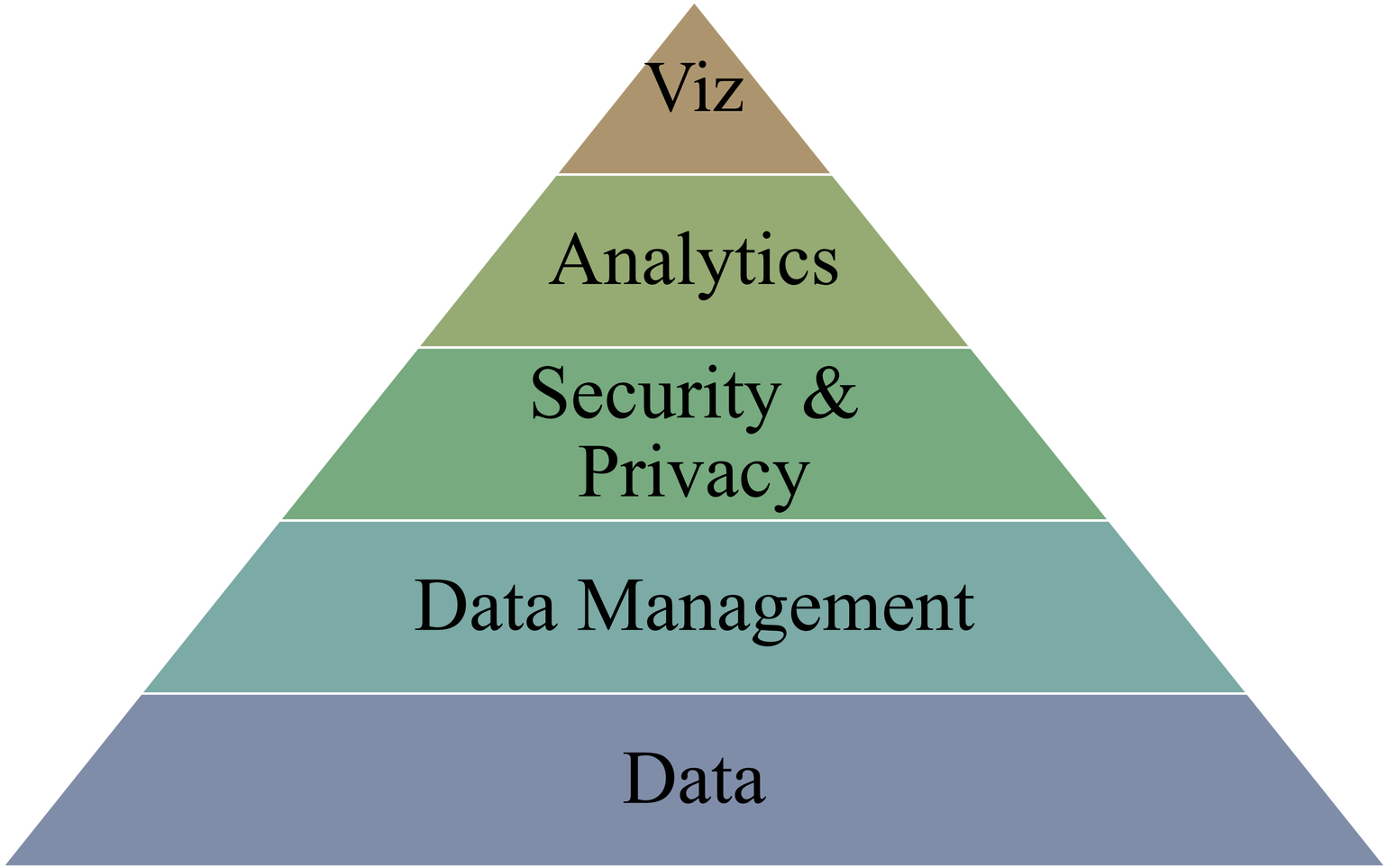
Process data are **noisy, complex**, and require detailed exploration

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



NAEP Process Data

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="2019-03-01T[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]

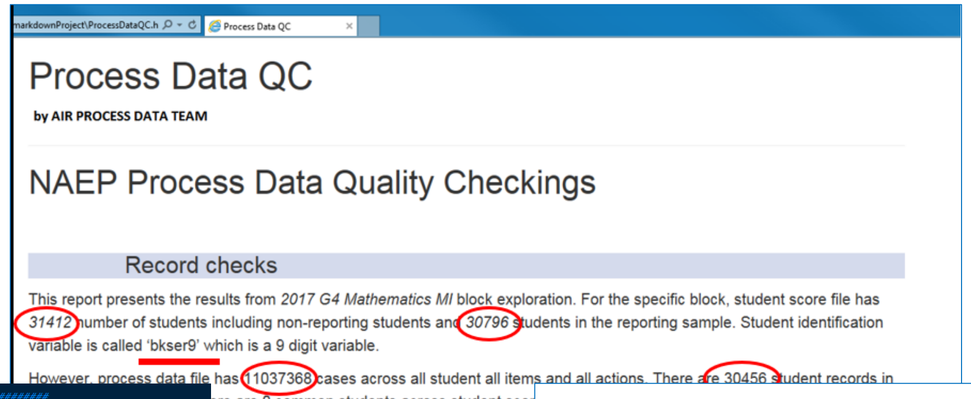
Event	Event	Event
Back	Eliminate Choice	Move Calculator
Calculator Buffer	Equation Editor Button	Next
Change Theme	Erase	Open Calculator
Clear Answer	Exit Item	Open Equation Editor
Clear Choice	Hide Timer	Receive Focus
Click Progress Navigator	Highlight	Scratchwork Draw Mode On
Close Calculator	Horizontal 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:

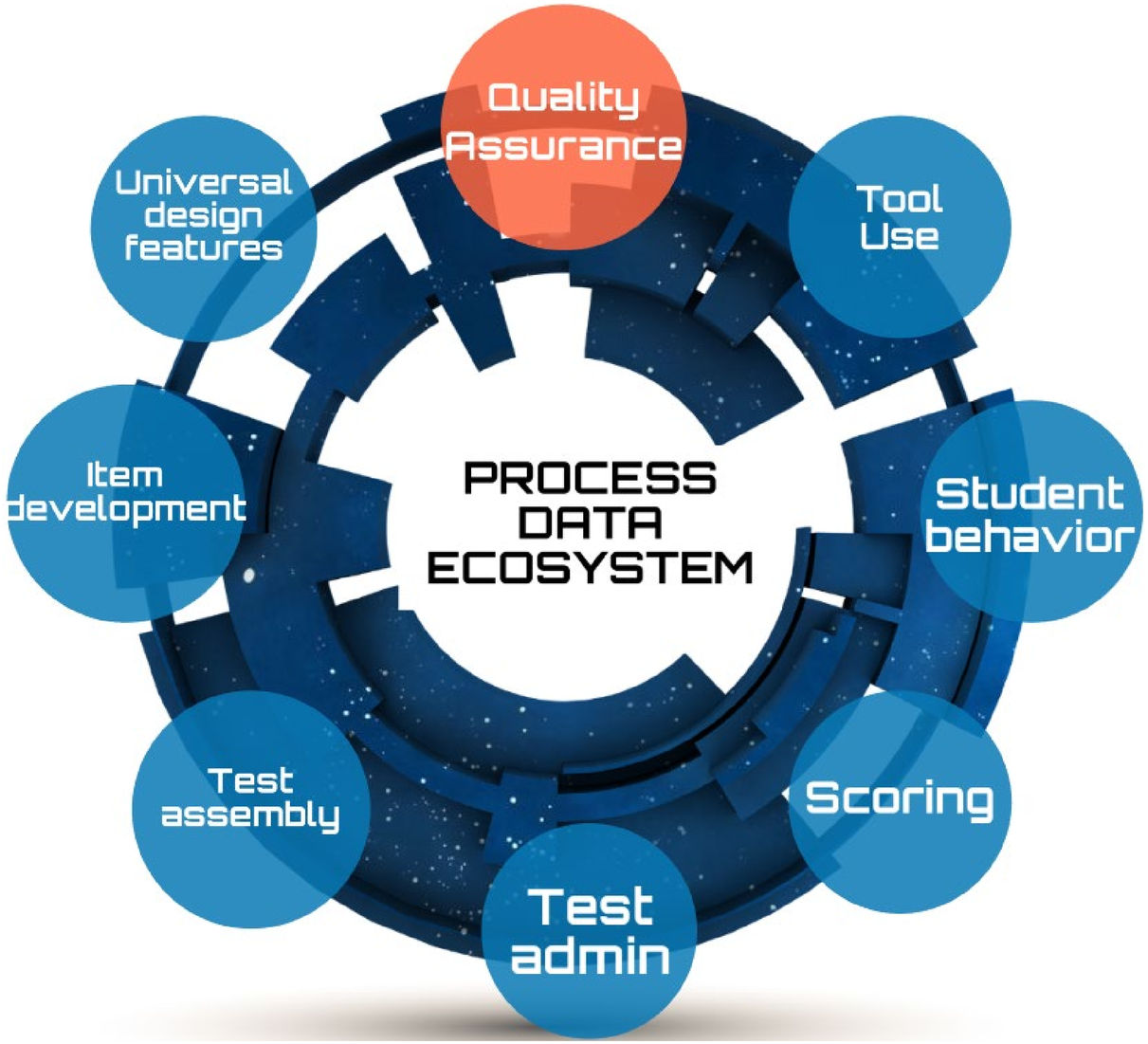
Data Request, Access, Storage, and Disposal Procedure

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
```

Gaining Insights: NAEP Process Data



Today's Examples

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

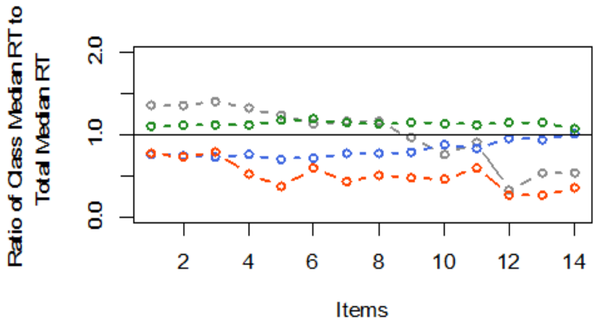
Students' Test Taking Behavior: Disengagement

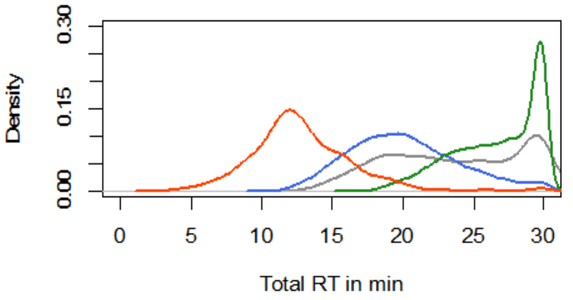
Our Research:

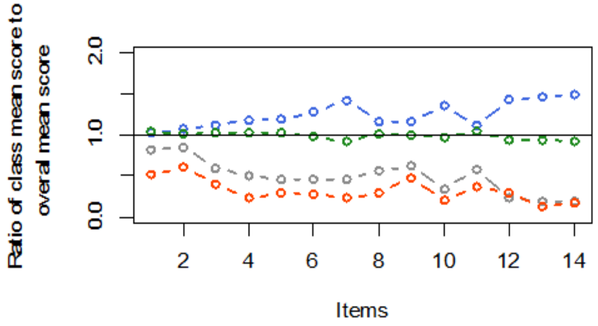
Can finite mixture modelling techniques be used on response time to identify distinct groups of students with different testing behaviors, such as disengagement and speededness?

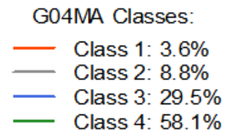
Key Findings:

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

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

Assessment Accommodations: Extended Time Accommodation

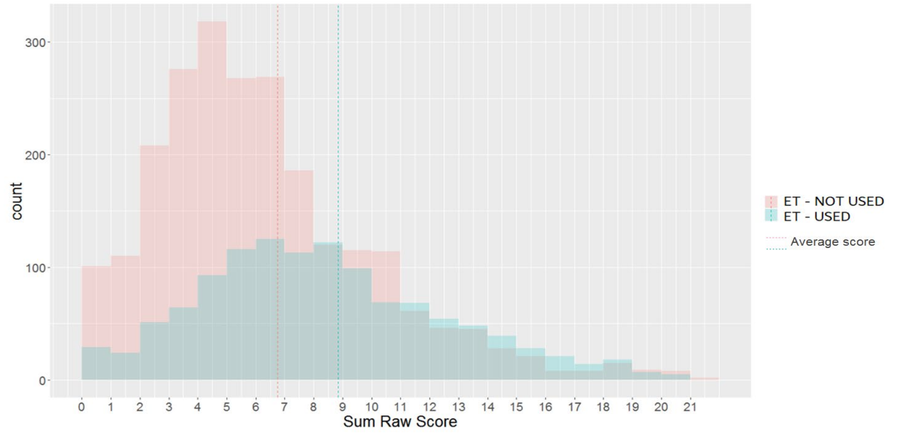
Our Research:

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



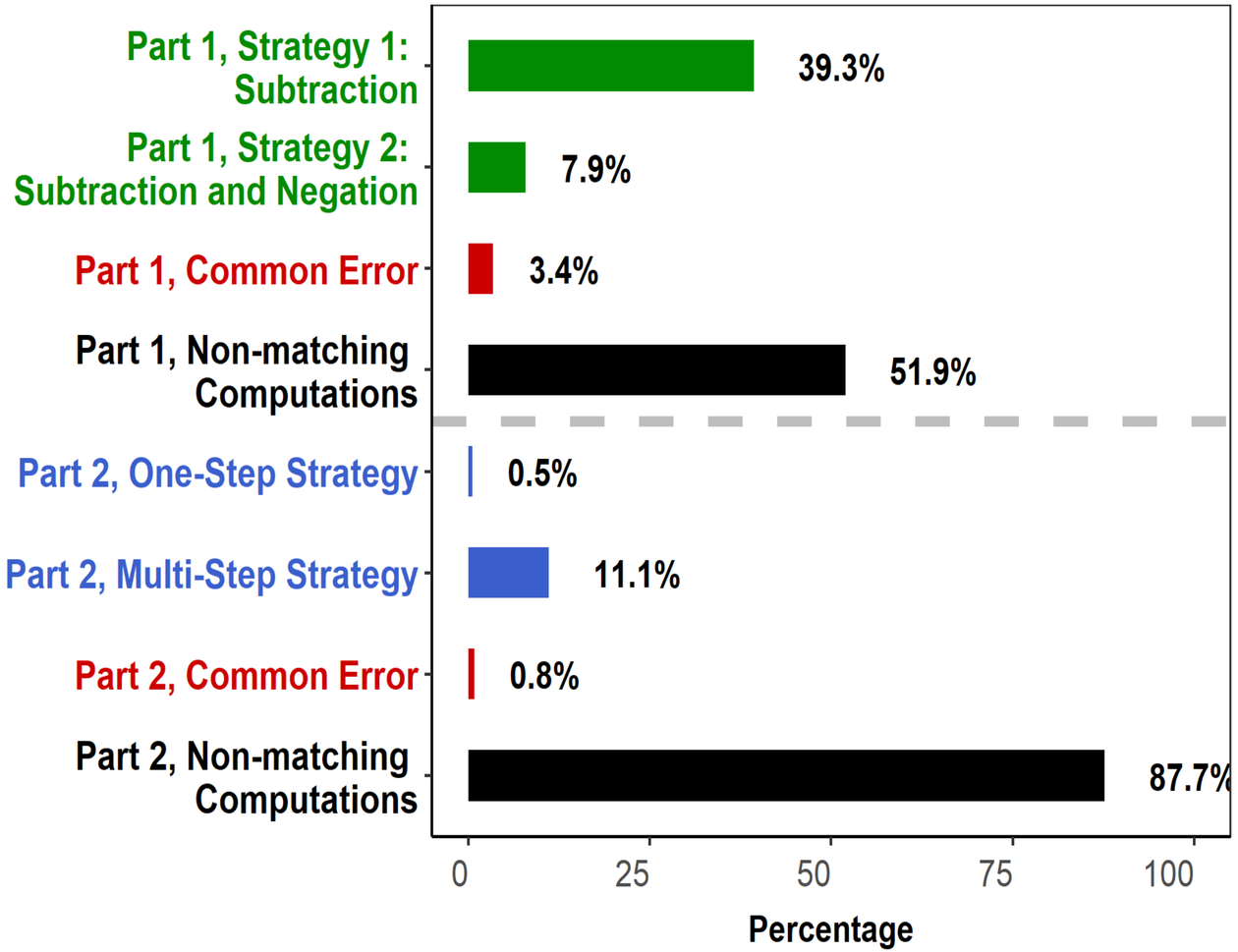
Assessment Features: Calculator Use Study

Our Research:

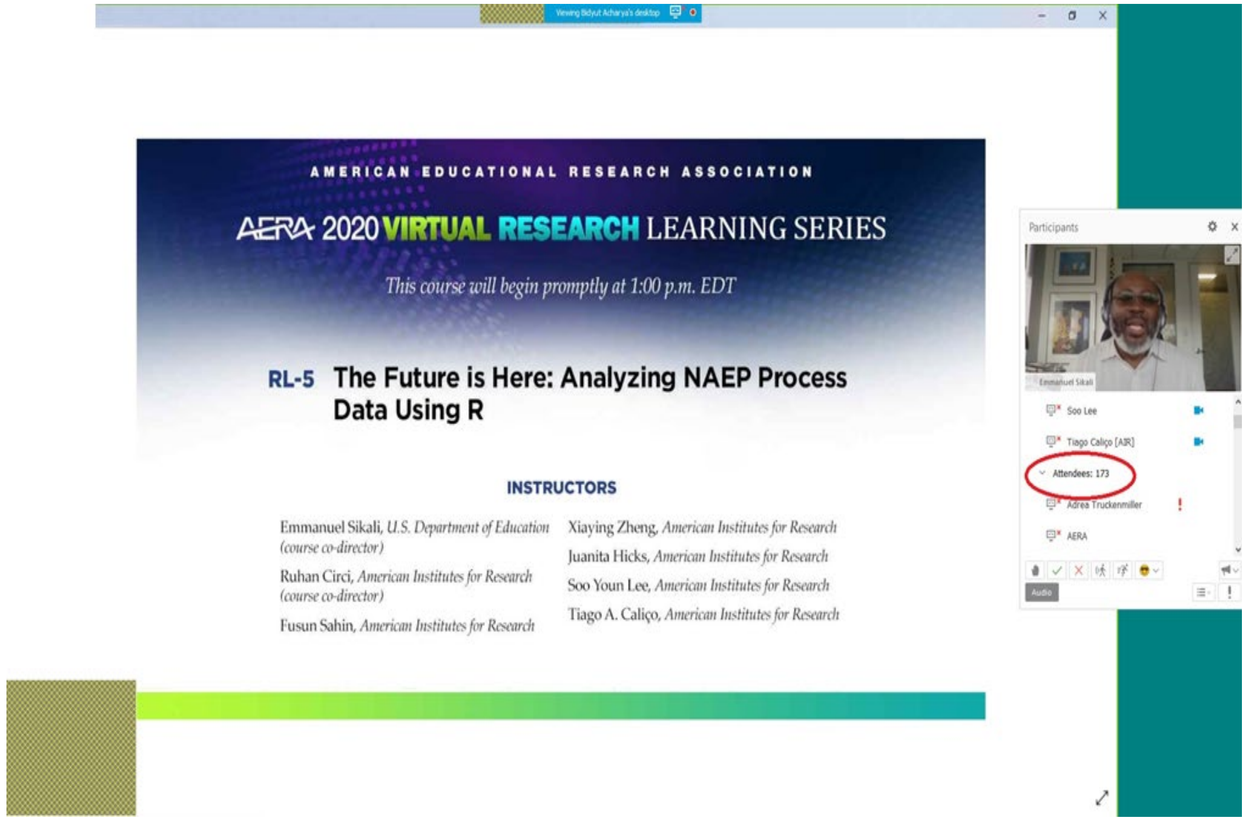
- What computations students do with the calculators?
- How often do students use computation strategies or commit common errors identified by experts using the calculator?

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.



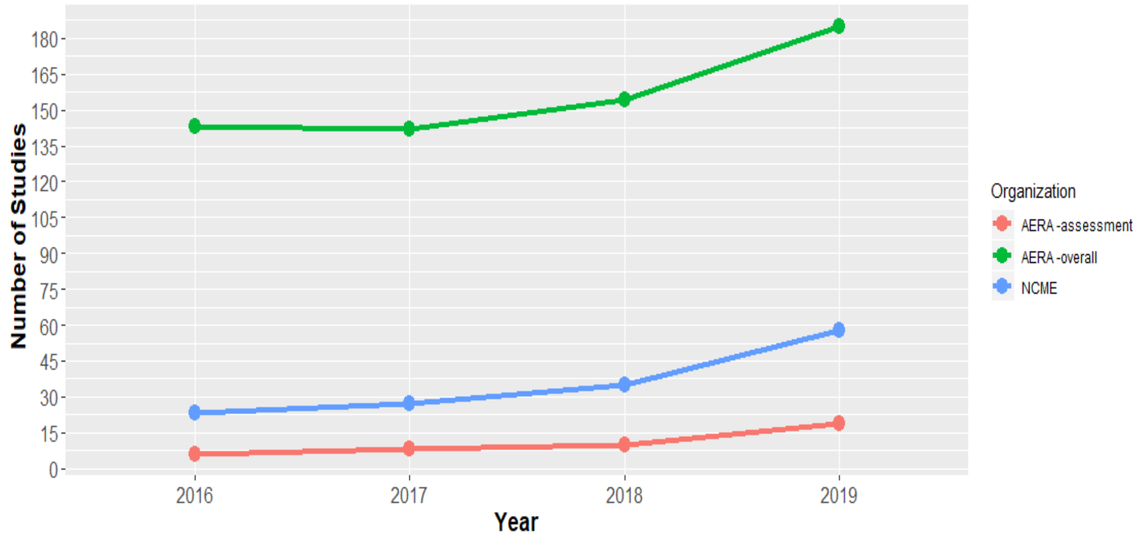
Process Data Community Building: 2019 NCME and 2020 AERA



Process Data Community Building: Special Interest Groups (SIGs)

AERA: Process Data in Digitally Based Assessments

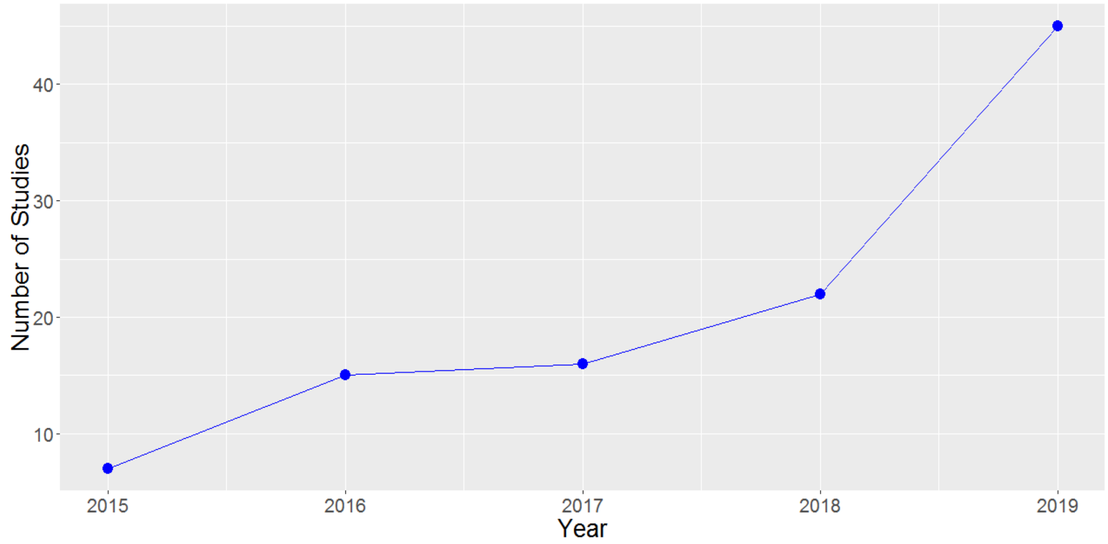
Process data research is an emerging topic that has sparked growing interest due to its novelty and yet unexplored potential



NCME: Big Data in Educational Measurement

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

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