



*Beyond results: Paving the way for the use of process data*  
*June 17-19<sup>th</sup>, 2020 (Virtual)*




# Advances and Challenges in Logging Documentation for Educational Assessments


Qiwei Britt He


Psychometrics, Statistics and Data Science  
Educational Testing Service



# Content

 What is logging documentation and why is it important?

 How logging documents are developed?

 What documents do we need?

 Challenges in logging documents

 Outlook and Recommendations



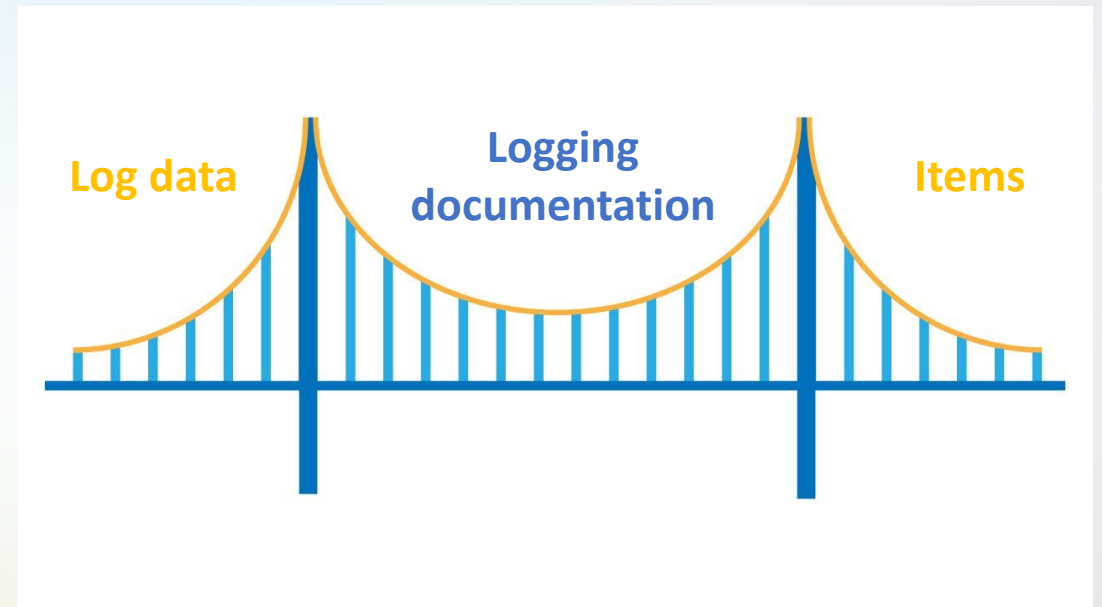
# What are log file and logging documentation

- A log (data) file, in a computing context, is the automatically produced and time-stamped documentation of events relevant to a particular system.
- Logging is the **act of keeping a log**. Documents that produce in the logging process are logging documentation.
- Logging documentation is a **KEY**.



# Why logging documentation is important?

- **A**ccessible document framework
- **B**ridge between raw log and item
- **C**omprehensive records
- **D**ynamic and meaningful interpretation
- **E**nsure standards and reproductivity



# Development of logging documentation

## Note on log file for PISA 2012 process data

### General notes:

SPSS files are provided for each item. In each of these files, the first four variables are the 3 letter country code, the 6 digit National Centre Code, the 7-digit school id, and the 5 digit student id. These fields can be used together to merge with the main database.

The fifth variable gives the id for the test form that the data come from.

The sixth variable is the type of event. This may be either system generated (start item, end item) or student generated (e.g. ACER\_EVENT, Click, Dbclick). The seventh variable is the event time, given in seconds from the beginning of the assessment. Note that in a small number of cases, the event times for a student's interactions are not sequential. This is an accurate representation of the sequence in which the events were written by the TAO system. The eighth variable is an event sequence number. Any remaining variables are event specific and they are described below.

Note that for a variety of reasons (for example if a student sat the assessment, but there were issues with their log files) a small number of student records have been set to "invalid".

### Item specific notes: Problem solving items

#### CP007

For all items in this unit, all click and double click events are included. Where the click event involves selecting or de-selecting a segment of the map, this is recorded in two rows, which have the same event time. In the first of the pair of rows, the state of the entire map is given, where a 1 in the sequence means that the segment was selected, and a 0, that it was not. The order of the map segments is shown through the example below. In this example, only the segment "nowhereSakharov" is selected.

#### Example:

```
Diamondnowhere"-0,"DiamondSilver"-0,"EmeraldLincoln"-0,"EmeraldUnity"-0,"LeeMandela"-0,"LincolnSato"-0,"MandelaEinstein"-0,"MarketLee"-0,"MarketPark"-0,"NobelLee"-0,"nowhereEinstein"-0,"nowhereEmerald"-0,"nowhereSakharov"-1,"nowhereUnity"-0,"ParkMandela"-0,"Parknowhere"-0,"SakharovMarket"-0,"SakharovNobel"-0,"Satonowhere"-0,"SilverMarket"-0,"Silvernowhere"-0,"UnityPark"-0,"UnitySato"-0
```

7 units released  
3-page general notes

## Codebook for PISA 2015 (FT) process data

### Codebook for Process Data in Released Science Items in PISA2015 Field Trial

#### Overview

[Codebook for response and background variables](#)

[General code for process data](#)

#### Standard units:

[Codebook for S600TraceEvent](#)

[Codebook for S613TraceEvent](#)

[Codebook for S639TraceEvent](#)

[Codebook for S644TraceEvent](#)

[Codebook for S655TraceEvent](#)

BeeColonyCollapseDisorder

FossilFuels

BluePowerPlant

VolcanicEruptions

GroundwaterExtractionAndEarthquakes

#### Interactive units:

[Codebook for S621TraceEvent](#)

[Codebook for S623TraceEvent](#)

[Codebook for S633TraceEvent](#)

AdjustableGlasses

RunningInHotWeather

EnergyEfficientHouse

He, Q. & Gonzalez, E. (2019). Codebook for process data and response data files for released science units in PISA 2015 Field Trial.

8 units released  
14 sheets codebook, 1462 entries

## Log Data Documentation (Framework) for PIAAC

https://piaac-logdata.tba-hosting.de/index.html

OECD Survey of Adult Skills (PIAAC) Log Data Documentation

This documentation provides an overview of process data recorded in log files in the PIAAC study.

Introduction

General Information on PIAAC log files and Booklet Orders

Please click on one of the items below to explore which event types can occur in an item. Use the NEXT button between pages of an item. You will be redirected to this list after every item by confirming "OK" to be proceed to the next task.

The following lists give examples for each response mode in Literacy and Numeracy. For Problem Solving in Technology-Rich Environments (PS-TRE), all items are displayed.

Literacy Items	Numeracy Items	Problem Solving Items
public	public	public
confidential	confidential	confidential
confidential	confidential	confidential

All domains log data released (with restriction), 10 web pages, 84 entries (aggregated level), include item information



# Developing a codebook (log document)

- Developing a codebook is time consuming but valuable and helpful.
- Codebook tells us the meaning of actions, structures and process.
- A codebook is generally built in a hierarchical way, including general codes across items and unique codes for a specific item.
- Codebook can be made in a fixed format (usually spreadsheet) or dynamic format with item information showing simultaneously.
- It is important to validate codebook with item developers and content experts to check whether the understanding of action meaning is accurate.

## Codebook for General Codes in Process Data in Released Science Items in PISA2015 Field Trial

### Note:

- (1) In this codebook sheet, the general codes in process data across different units are listed, mainly including **column names, event names, module, target and datatype**.
- (2) The item type (e.g., multiple choice items, checkbox items) have general codes in process data, which are listed in the last table in this sheet. These codes are recorded in "datavalue" column in each item.
- (3) Specific codes for actions/values related to items are list in codebooks by each unit.
- (4) Process data usually have noise and need to be cleaned before they can be used for analysis. The contents (actions) in datavalue column that occur less than 5 times in the sample are discarded from the codebook.

### Column names

Column Names	Meaning
CNSTID	Matching key ID between process data (TraceEvent file) and response data (BQ)
country	Synthetic Country, cn1, cn2, cn3
unitId	unit name in PISA2015: SCI/S600-BeeColonyCollapseDisorder
itemId	Item ID in PISA 2015 process data, need to match with the response scored item ID in BQ
itemName	Item name (page) within the unit
eventCounter	Sequence of event, labeled from 0, end with the last event within the unit
time	Timestamp, unit as milliseconds
event_name	Contents of the event, usually needs to be interpreted together with columns of "modeuleId" and "target".
moduleId	The module of the event, e.g., platform, scoring or controller.
target	Actions, usually needs to be interpreted together with columns of "datavalue".
datatype	Variables embedded in process data, e.g., total number of actions, total response time, and etc.
datavalue	Value of datatype
VAR_AllRowsCorrect	The number of rows are correctly labeled to support the interpretations
VAR_DataTable	The eventual data table created by test takers (record per each row)
VAR_NumRows	Number of rows recorded in table

## Codebook for S613\_TraceEvent\_3cnty\_synthetic.csv

### Note:

- (1) Action sequences in standard units were not completely captured. For instance, keypress actions were captured for some students, but not all. However, in the variable "number of actions", the keypress actions were counted.
- (2) Some test takers do have "keypress" actions, but the contents of keypress were not captured. It means we only know the test takers make input in the open-ended response, but we do not know what they typed into by keypress each time.

### ItemID in TraceEvent Match with Response Data and Sequence Page

Item ID in Process Data	(Scored) Item ID in Response Data	itemName	Item type
CS613Q01	CS613Q01S	item 1	multiple choice (one option)
CS613Q02	DS613Q02C	item 2	open-ended (2 subitems)
CS613Q03	DS613Q03C	item 3	open-ended

Note: Each item has its own page

### Contents in "datavalue" Column

datavalue	meaning
clockDiv	click on the clock part on the top of the page
container	click on the type in box
diagram2	click on the diagram in CS613Q03
FossilFuels1	click on text passage in CS613Q02
FossilFuels2	click on text passage in CS613Q03
graph1	click on graph in CS613Q02
header	click on the header
headerTitle	click on the header title area
help	click on the help function (question mark button)
legend2	click on diagram legend in CS613Q03
navigation	click in the navigation arrows (back or next)



# Developing a codebook – two approaches

## Item interface approach

**Running in Hot Weather**  
Question 1 / 5

**How to Run the Simulation**

Run the simulation to collect data based on the information below. Select from the drop-down menus to answer the question.

A runner runs for one hour on a hot, dry day (air temperature 40°C, air humidity of 20%). The runner does not drink any water.

What health danger does the runner encounter by running under these conditions?

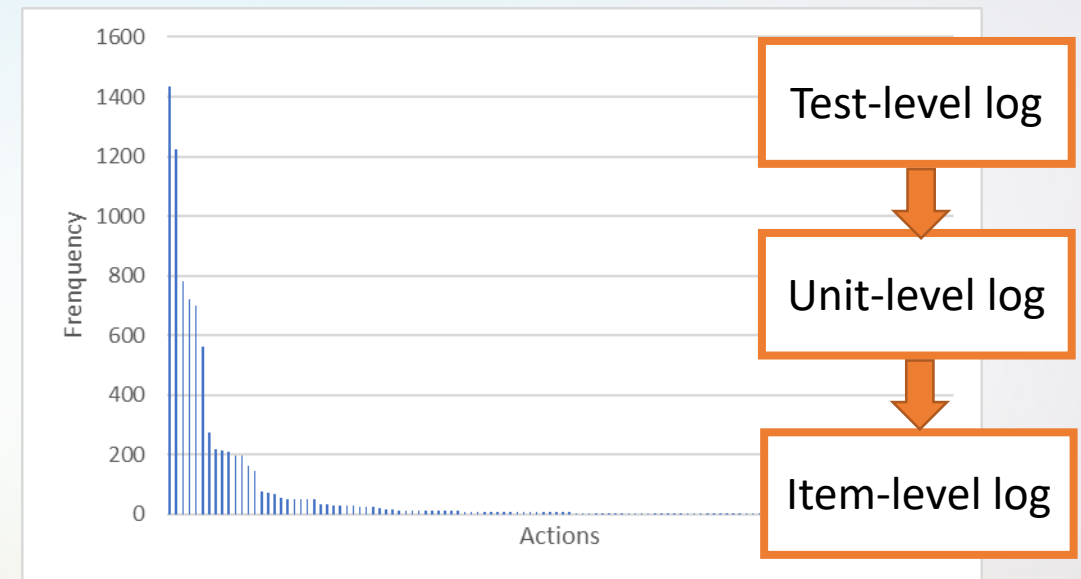
The health danger that the runner encounters is .

This is shown by the  of the runner after a one-hour run.

Air Temperature (°C)	Air Humidity (%)	Drinking Water	Sweat Volume (Litres)	Water Loss (%)	Body Temperature (°C)
25	20	Yes	1.0	0.0	39.0



## Data driven approach

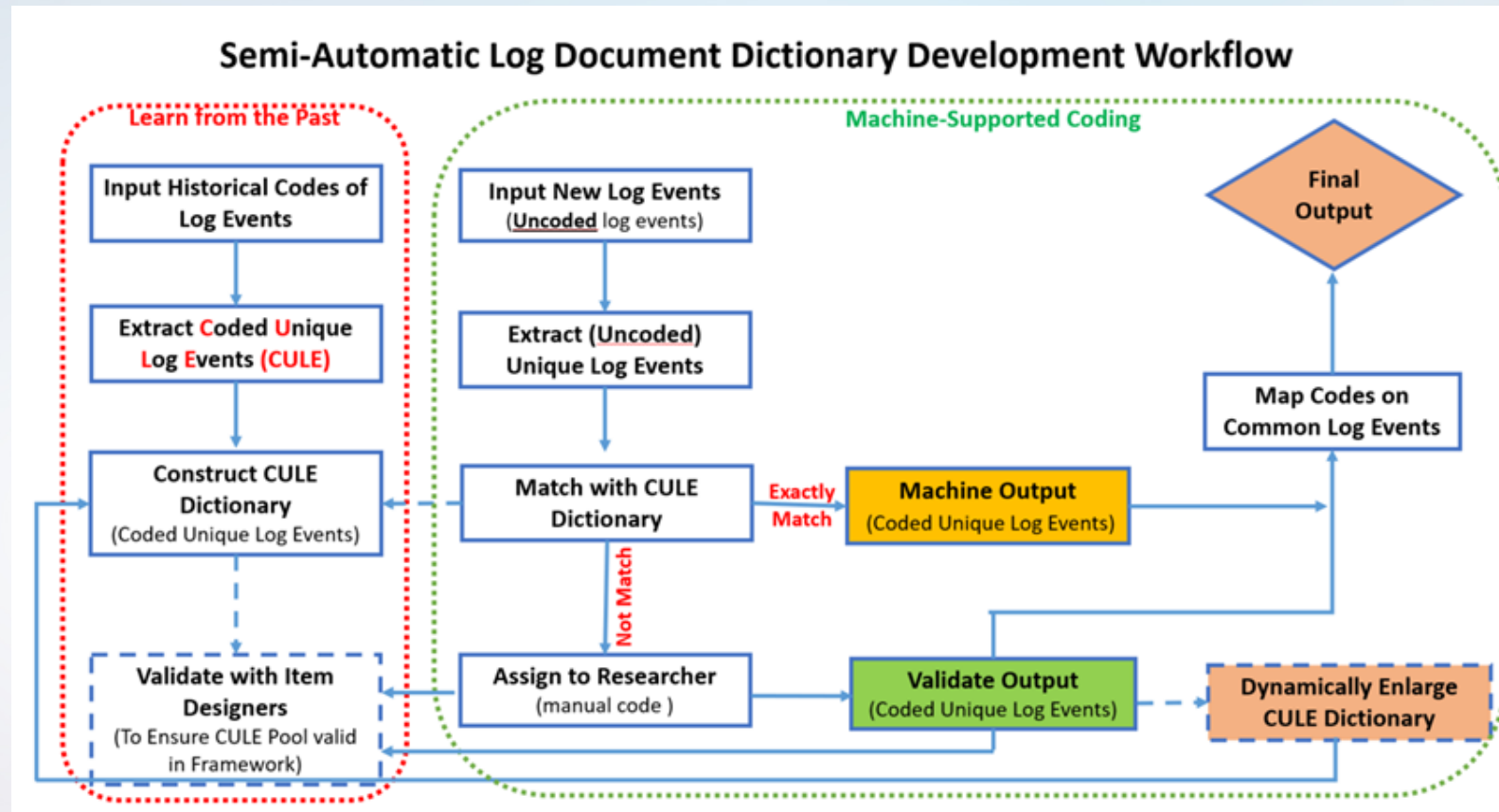


- Familiar with item interface developers
- Interpretable clicks, clear structures
- Missing actions from keyboard
- Over-efforts on actions that may never be used

- Multilevel logging documentation, may also support item/data release by different levels
- Principle: higher level to lower level, action with higher frequency to lower frequency



# Machine-supported log document dictionary



Reference: Yamamoto, He, Shin, & von Davier, M. (2018).

- Combine manual and auto efforts in building up a dictionary
- Requirement on standard raw log data coding
- Sets of dictionary could be established for different program



# A joint document with log events and item

PISA 2015

**Running in Hot Weather**  
Question 3 / 5

How to Run the Simulation

Run the simulation to collect data based on the information below. Click on a choice, select data in the table, and then type an explanation to answer the question.

When the air humidity is 60%, what is the effect of an increase in air temperature on sweat volume after a one-hour run?

Sweat volume increases  
 Sweat volume decreases

Select two rows of data in the table to support your answer.

What is the biological reason for this effect?

The simulation interface includes several gauges: Sweat Volume (Litres) from 0 to 3, Water Loss (%) from 0 to 5, and Body Temperature (°C) from 16 to 42. A 'Run' button is visible. The data table below shows the following data points:

Air Temperature (°C)	Air Humidity (%)	Drinking Water	Sweat Volume (Litres)	Water Loss (%)	Body Temperature (°C)
30	40	No	1.2	1.8	39.3
30	20	Yes	1.1	0.0	39.1

	CNSTID	country	unitId	itemId	itemName	eventCour	time	event_nam	moduleId	target	datatype	datavalue
	761	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	36	1398755338929	QuestionL	platform	MODULE	
	762	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	37	1398755338930	stimulus1_13	platform	MODULE	
	763	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	38	1398755394065	click	td	id	dtTROTD1
	764	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	38	1398755394065	click	td	timeStamp	1.39876E+12
	766	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	39	1398755395744	click	span	id	stimulus1_13
	767	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	39	1398755395744	click	span	timeStamp	1.39876E+12
	770	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	40	1398755409977	click	div	timeStamp	1.39876E+12
	772	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	41	1398755417576	click	span	id	stimulus1_44
	773	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	41	1398755417576	click	span	timeStamp	1.39876E+12
	775	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	42	1398755417580	click	input	id	drinking-no
	776	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	42	1398755417580	click	input	timeStamp	1.39876E+12
	778	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	43	1398755417582	change	input	id	drinking-no
	779	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	43	1398755417582	change	input	timeStamp	1.39876E+12
	780	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	44	1398755418811	click	button	id	btn-run
	781	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	44	1398755418811	click	button	timeStamp	1.39876E+12
	783	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	45	1398755420992	click	span	id	stimulus1_42
	784	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	45	1398755420992	click	span	timeStamp	1.39876E+12
	786	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	46	1398755420997	click	input	id	drinking-yes
	787	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	46	1398755420997	click	input	timeStamp	1.39876E+12
	789	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	47	1398755420999	change	input	id	drinking-yes
	790	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	47	1398755420999	change	input	timeStamp	1.39876E+12
	791	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	48	1398755422017	click	button	id	btn-run
	792	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	48	1398755422017	click	button	timeStamp	1.39876E+12
	794	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	49	1398755426973	click	td	id	dtTROTD2
	795	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	49	1398755426973	click	td	timeStamp	1.39876E+12
	797	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	50	1398755427815	click	td	id	dtTR1TD3



# Developing a dynamic codebook

Unit 2  
You would like to copy some music files to your portable music player.  
The music player has room for 20 MB and you want as many files as possible. You want to include only jazz and rock music.  
Select the files to include.  
Once you have selected the files, click Next to continue.

Title	Size	Time	Artist	Genre
<input type="checkbox"/> A Foreign Affair	14.8 MB	11:40	Don Rader Quartet	Jazz
<input type="checkbox"/> About the Blues	4.3 MB	3:08	Julie London	Blues
<input type="checkbox"/> Another Mind	7.8 MB		Romi Uehara	Jazz
<input type="checkbox"/> Blue Trane	10 MB		John Coltrane	Jazz
<input type="checkbox"/> Don't Give up on Me	3.5 MB		Olomon Burke	Blues
<input type="checkbox"/> Far Out	5.3 MB		Antonio Farao	Jazz
<input type="checkbox"/> Fire and Water	5.3 MB		Lee	Blues
<input type="checkbox"/> If	4.9 MB		Yriam Alter	Jazz
<input type="checkbox"/> Imagine	2.2 MB		John Lennon	Rock
<input type="checkbox"/> Inclined	7.1 MB		Carol Welsman	Jazz
<input type="checkbox"/> On an Island	16 MB		David Gilmore	Blues
<input type="checkbox"/> Pass It On	3.1 MB		Bert Calvo	Jazz
<input type="checkbox"/> Raindrops, Raindrops	5.2 MB		Marin Krog	Jazz
<input type="checkbox"/> Say You Will	8.8 MB		Beatwood Mac	Rock
<input type="checkbox"/> Skin Deep	7.1 MB		Buddy Guy	Blues
<input type="checkbox"/> Speak No Evil	6.9 MB		Lora Purim	Jazz
<input type="checkbox"/> The Other Side of Blue	6.5 MB		San Shy & Jobo	Jazz
<input type="checkbox"/> The Rise	7.3 MB		Jillien Lourau	Jazz
<input type="checkbox"/> The Rising	4.5 MB	4:50		

Sort dialog box options:  
Sort by: Choose a column title  
Then by: Choose a column title  
And then by: Choose a column title  
Buttons: OK, Cancel

```
<taoEvent Name="stimulus" Type="BUTTON" Time="12345">
  id=sortCancel
</taoEvent>
<taoEvent Name="stimulus" Type="DOACTION" Time="12345">
  action=as://closeWindow(sortwindow)
</taoEvent>
```

Total Size Selected (MB) [ ]

Cited from Goldhammer, F., & Hahnel, C. (2019). PIAAC log files and process data: what can the process data tell us about the process of test taking, workshop held by American Institutes for Research, Arlington, VA, USA.

Bee Colony Collapse Disorder  
Question 2 / 5

Refer to "Bee Colony Collapse Disorder" on the right. Type your answer to the question.

Understanding colony collapse disorder is important for people who keep and study bees, but colony collapse disorder also has an effect beyond the bees. People who study birds have identified an impact. The sunflower is a food source for both bees and certain birds. Bees feed on the nectar of the sunflower, while the birds feed on the seeds.

Given this relationship, why might the disappearance of bees result in a decline in the bird population?

**BEE COLONY COLLAPSE DISORDER  
Exposure to Imidacloprid**

Scientists believe that there are multiple causes for colony collapse disorder. One possible cause is the insecticide imidacloprid, which may cause bees to lose their sense of orientation when outside the hive.

Researchers tested whether exposure to imidacloprid leads to colony collapse. In a number of hives, they added the insecticide to the bees' food for three weeks. Different hives were exposed to different concentrations of the insecticide, measured in micrograms of insecticide per kilogram of food ( $\mu\text{g}/\text{kg}$ ). Some hives were not exposed to any insecticide.

None of the colonies collapsed immediately after exposure to the insecticide. However, by week 14, some of the hives had been abandoned. The following graph records the observed results:

Percentage of Colonies Collapsed

Number of Weeks After Exposure to Insecticide

Weeks	0 $\mu\text{g}/\text{kg}$	20 $\mu\text{g}/\text{kg}$	400 $\mu\text{g}/\text{kg}$
10	0%	0%	0%
12	0%	0%	0%
14	0%	25%	50%
16	0%	25%	50%
18	0%	25%	100%
20	25%	75%	100%
22	25%	100%	100%

Cited from He, Q. & Gonzalez, E. (2019). Opportunity versus Challenge: Exploring usage of log-file and process data in international large scale assessments, workshop held by Educational Research Center of Ireland and ETS, Dublin, Ireland.



# What documents do we need in the framework?

## Log files

- Raw log files
- Structuralized datasets (timestamped event logs)

## Codebook

- Log-item codebook (by multilevel)
- (Dynamic) codebook

## Item files

- Item description
- Scoring rules
- Interface
- Item calibrations and parameters
- Item design / optimal problem-solving strategies
- Item sequence and position

## Respondent

- Response data (credits)
- Sample description
- Background information

## Features

- Aggregated level and item level
- Sequence-based features (n-grams, action sequence, etc.)
- Cognitive features
- Timing feature

## Data files

- Response data
- Process data
- Merged data



# Challenges in logging documents

- Big sized log data is challenging for storage and developing log documentation. Manual work has to be involved in interpretation of log events and validation.
- Log data is still as a “byproduct” in most educational assessments. Missing values in log events result in incomplete records in documentation.
- Confidentiality issue may also impact logging documents. Published data and documentation may not be matched.

PISA 2015

**Adjustable Glasses**  
Question 3 / 5

▶ **How to Run the Simulation**

Run the simulation to collect data based on the information below. Select from the drop-down menu to answer the question.




Anna sees both near and distant objects in focus.  
How do adjustments to the glasses affect Anna's vision?

Adding fluid to the lens makes  objects appear out of focus.  
Removing fluid from the lens makes  objects appear out of focus.

**Amount of Fluid in Lens** **Distance from Tree**

Amount of Fluid in Lens: -2 -1 0 1 2  
Distance from Tree: near midway distant

**Run**

		Amount of Fluid in Lens				
		-2	-1	0	+1	+2
Distance from Tree	Near					
	Midway					
	Distant					



# Recommendations for future work

- To leverage efforts from multiple perspectives (e.g., item designers, researchers, content experts and etc.) in frontline to predefine information (hypothesis, features, strategies, etc.) that needs to be captured before the item is developed.
- To publish log data, log document and item interface (preferred) together. Unmatched information may cause less meaningful results.
- To standardize coding of raw log data to enhance the proportion of automation in logging documents development.
- To keep log documentation maintenance, benefit for open science and reproductivity.



# Reference

- Goldhammer, F., Naumann, J., Selter, A., Toth, K., Rolke, H., and Klieme, E. (2014). The time on task effect in reading and problem-solving is moderated by task difficulty and skill: insights from a computer-based large-scale assessment. *J. Educ. Psychol.* 106, 608–626. doi: 10.1037/a0034716
- Greiff, S., Wustenberg, S., and Avvisati, F. (2015). Computer-generated log-file analyses as a window into students' minds? A showcase study based on the PISA 2012 assessment of problem-solving. *Comput. Educ.* 91, 92–105. doi: 10.1016/j.compedu.2015.10.018
- Han, Z., He, Q., & von Davier, M. (2019). Predictive Feature Generation and Selection Using Process Data from PISA Interactive Problem-Solving Items: An Application of Random Forests. *Frontiers in Psychology*, 10: 1421.
- He, Q., Borgonovi F. & M. Paccagnella, M. (2019). Using process data to understand adults' problem-solving behaviour in the Programme for the International Assessment of Adult Competencies (PIAAC): Identifying generalised patterns across multiple tasks with sequence mining. *OECD Education Working Papers*, No. 205, OECD Publishing, Paris, <https://doi.org/10.1787/650918f2-en>.
- He, Q., & von Davier, M. (2016). Analyzing Process Data from Problem-Solving Items with N-Grams: Insights from a Computer-Based Large-Scale Assessment. In Y. Rosen, S. Ferrara, & M. Mosharraf (Eds.) *Handbook of Research on Technology Tools for Real-World Skill Development* (pp. 749-776). Hershey, PA: Information Science Reference.
- Lee, Y. H., and Haberman, S. J. (2016). Investigating test-taking behaviors using timing and process data. *Int. J. Test.* 16, 240–267. doi: 10.1080/15305058.2015.1085385
- Tang, X., Wang, Z., He, Q., Liu, J., & Ying, Z. (2019). Latent feature extraction for process data via multidimensional scaling. *arXiv preprint arXiv:1904.09699*.
- Yamamoto, K., He, Q., Shin, H. J., & von Davier, M. (2018). Development and implementation of a machine-supported coding system for constructed-response items in PISA. *Psychological Test and Assessment Modeling*, 60(2), 145–164.

- A special issue of “Exploring Usage of Log-File and Process Data in International Large-Scale Assessments” in *Large-Scale assessments in Education* will be published in October.

Guest editors:

*Qiwei (Britt) He (ETS)*

*Caroline McKeown (ERC)*

*Jude Cosgrove (ERC)*

*Irwin Kirsch (ETS)*





# Thank you very much!

Qiwei Britt He  
qhe@ets.org

