Beyond results: Paving the way for the use of process data
June 17-19th, 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

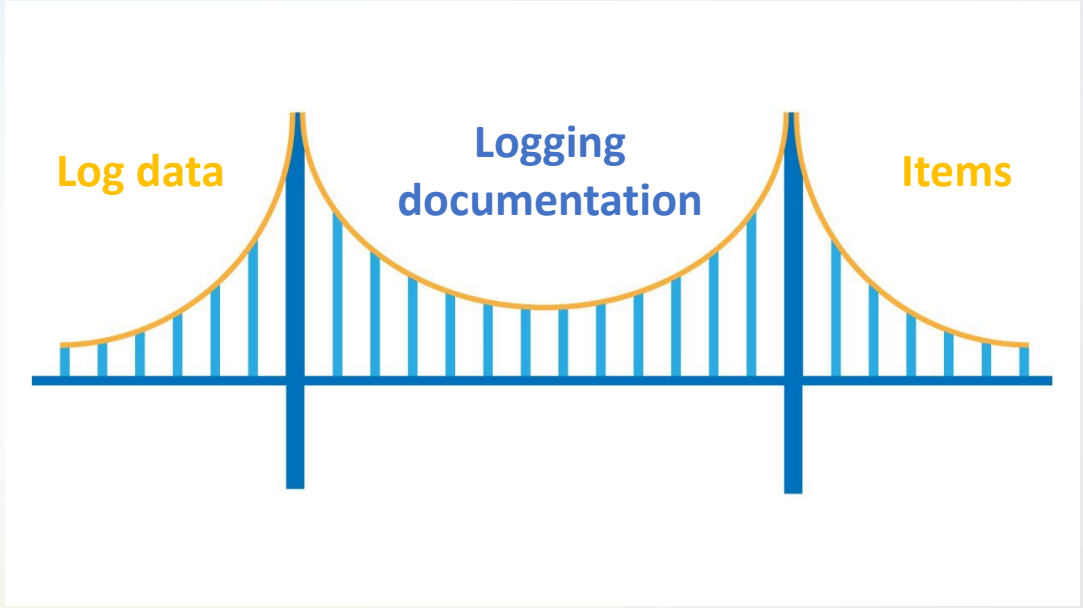
Outlook and Recommendations

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**.

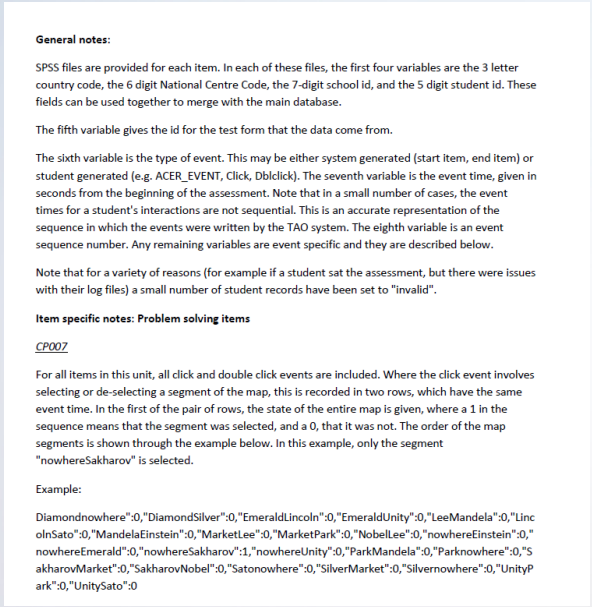
Why logging documentation is important?

- **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, Dblclick). 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	BeeColonyCollapseDisorder
Codebook for S613TraceEvent	FossilFuels
Codebook for S639TraceEvent	BluePowerPlant
Codebook for S644TraceEvent	VolcanicEruptions
Codebook for S655TraceEvent	GroundwaterExtractionAndEarthquakes

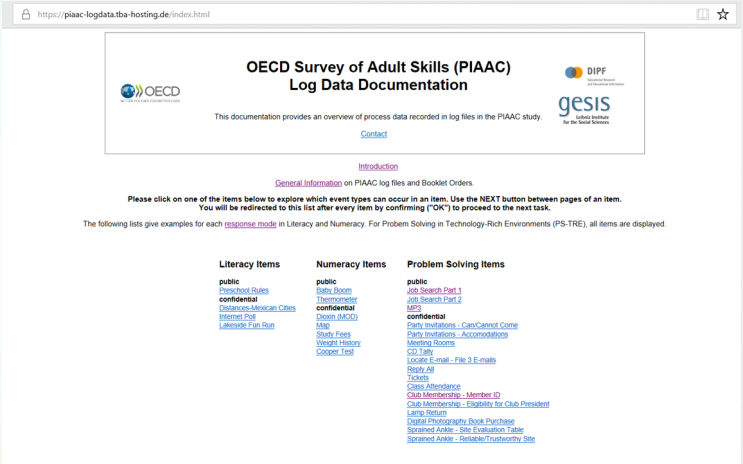
Interactive units:

Codebook for S621TraceEvent	AdjustableGlasses
Codebook for S623TraceEvent	RunningInHotWeather
Codebook for S633TraceEvent	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



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 miliseconds
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 controler.
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 "numer 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	mutiple choice (one option)
CS613Q02	DS613Q02C	item 2	open-ended (2 subitems)
CS613Q03	DS613Q03C	item 3	open-ended

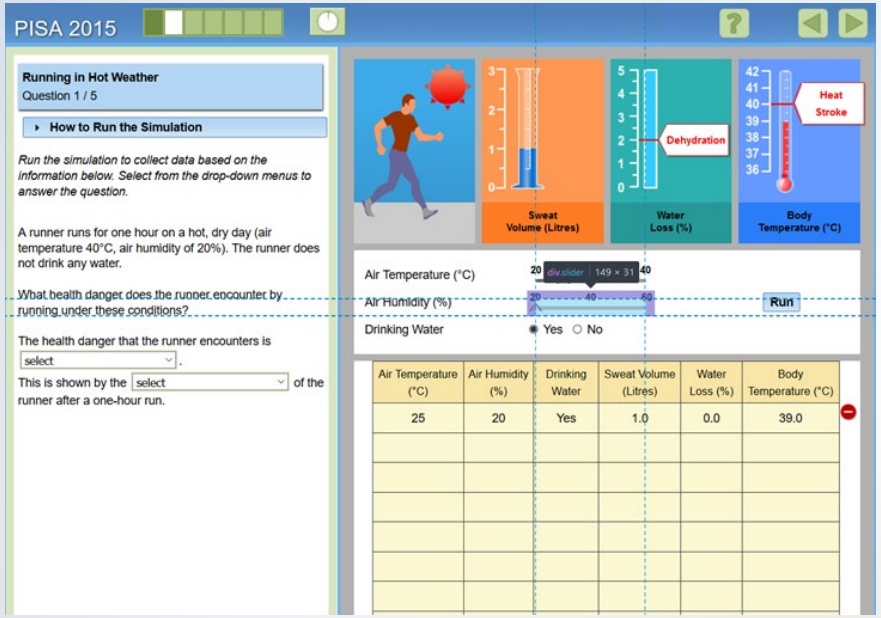
Note: Each item has its own page

Contents in "datavalue" Column

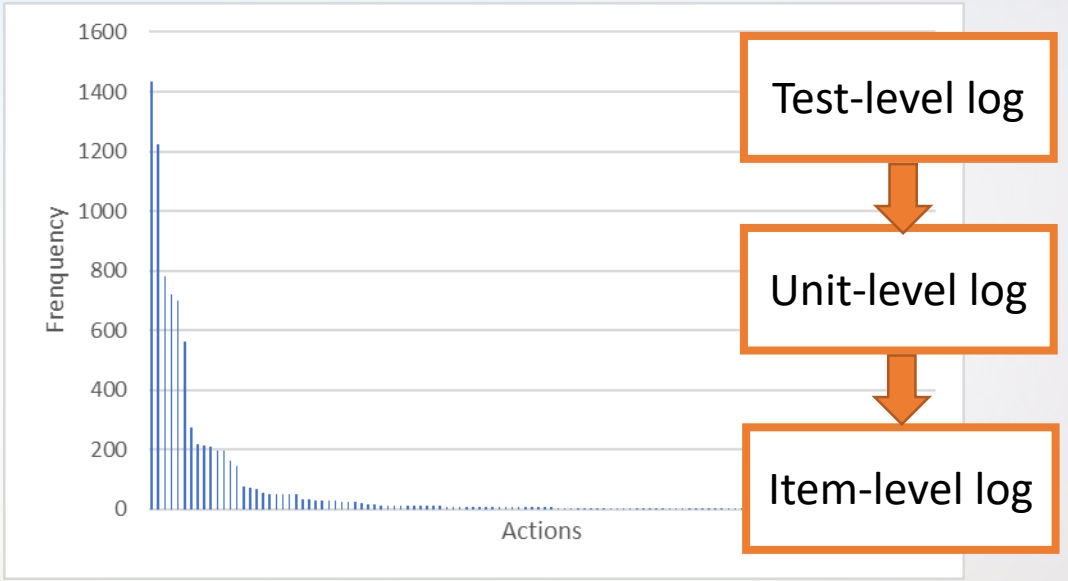
datavalue	meaning
clockDiv	click on the clock part on the top of the page
container	click on the type in box
diagram2	click on the diagram in CS613Q03
FossilFuels1	click on text passage in CS613Q02
FossilFuels2	click on text passage in CS613Q03
graph1	click on graph in CS613Q02
header	click on the header
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



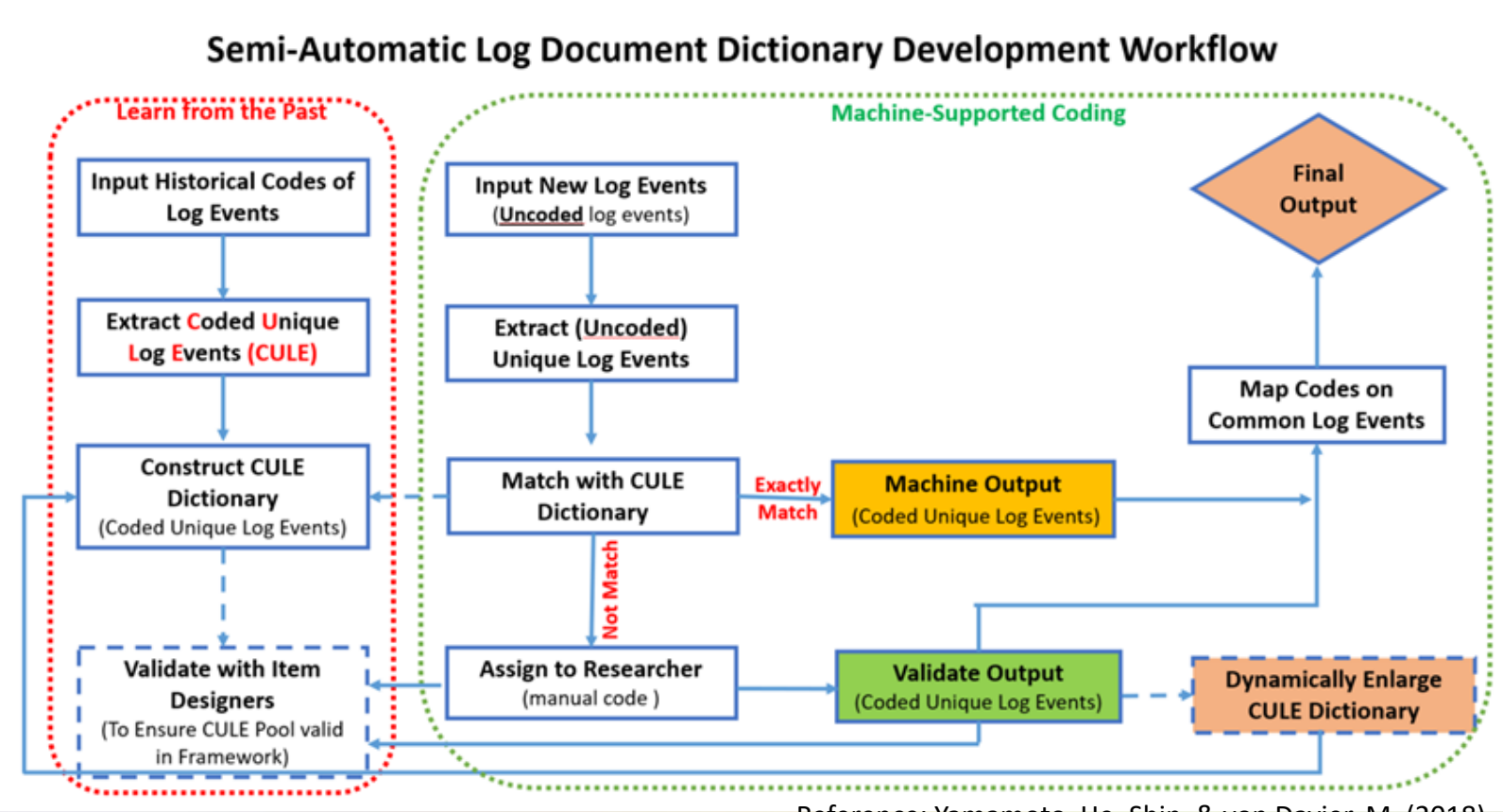
Data driven approach



- 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

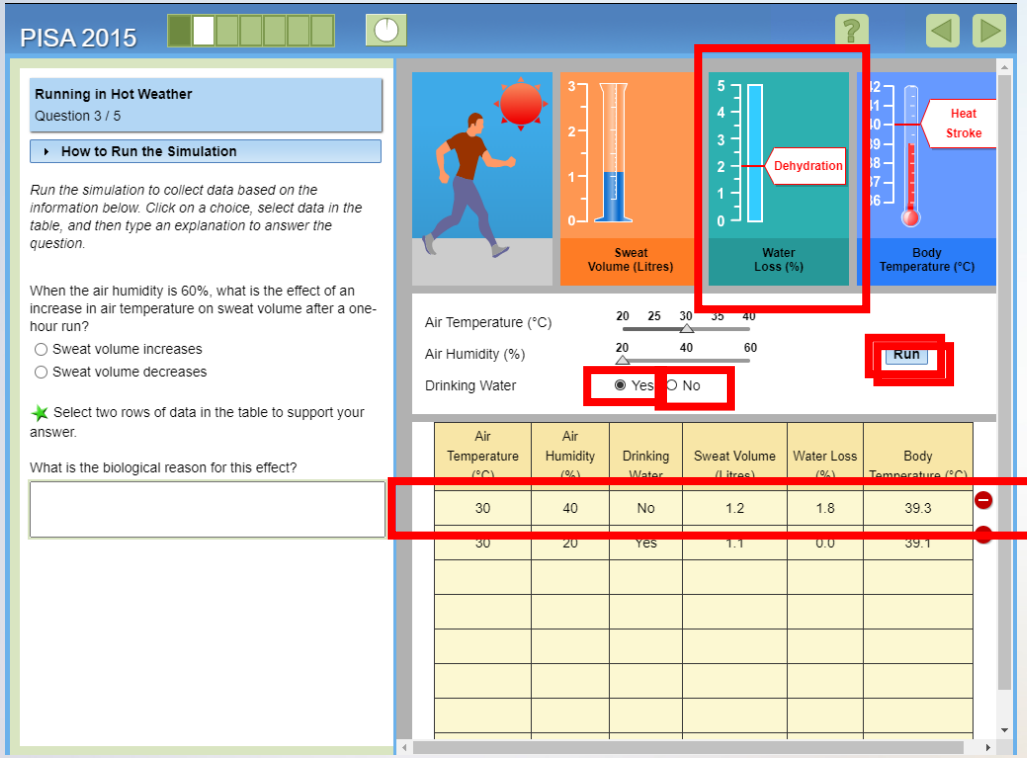


Reference: Yamamoto, He, Shin, & von Davier, M. (2018).

- 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

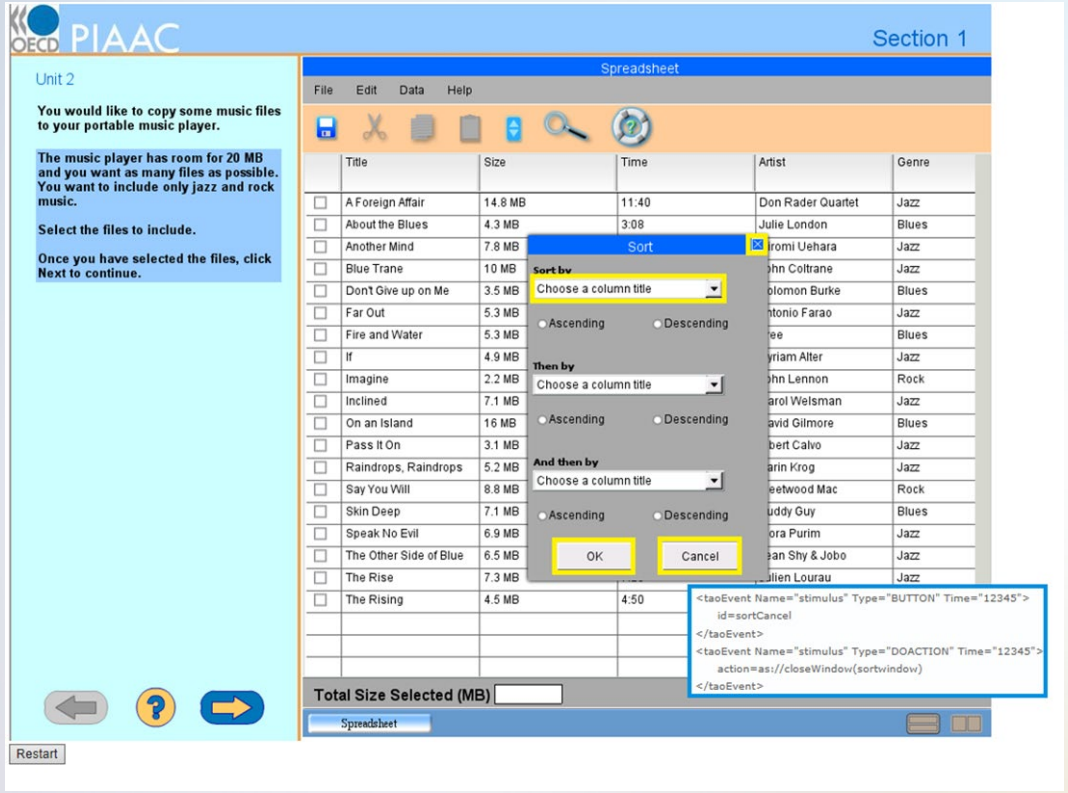
Citation: Yamamoto, K., He, Q., Shin, H. J., & von Davier, M. (2018). Development and implementation of a machine-supported coding system for constructed-response items in PISA. Psychological Test and Assessment Modeling, 60(2), 145–164.

A joint document with log events and item

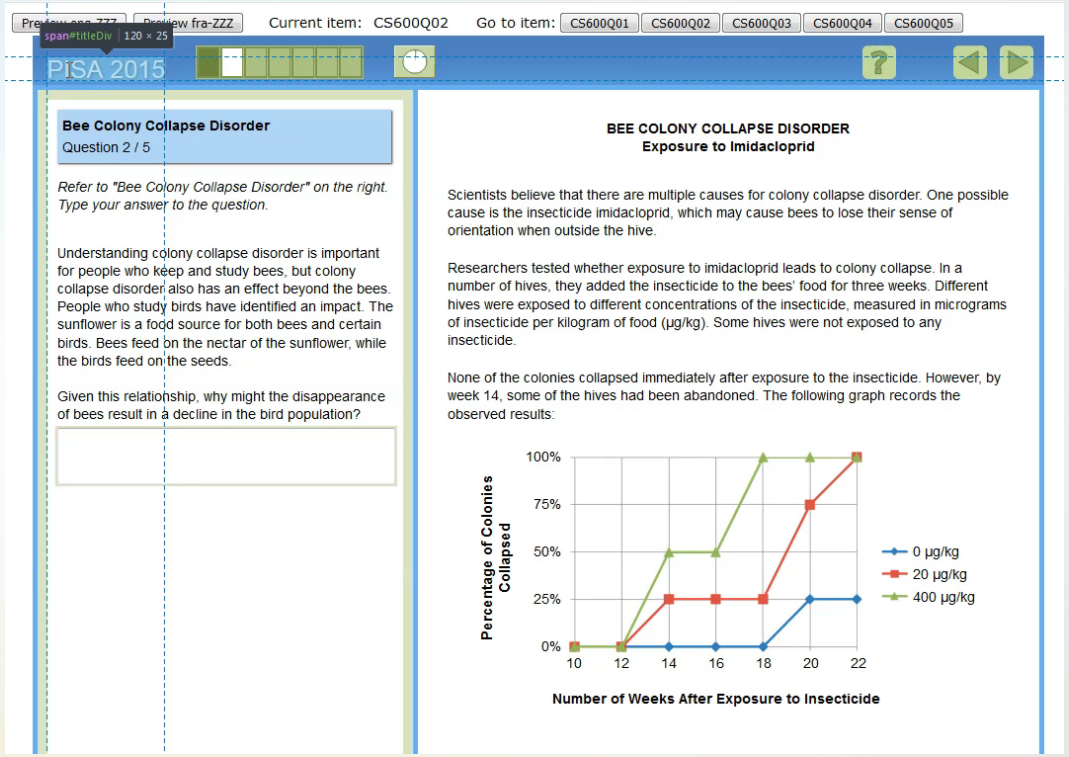


	CNStID	country	unitId	itemId	itemName	eventCour	time	event_nan	moduleId	target	datatype	datavalue
761	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	36	1398755338929	QuestionL	platform	MODULE		
763	cn1-00004	cn1	SCI/S623-RunningInHotWeather	CS623Q02	item3	38	1398755394065	click		td	id	dtTR0TD1
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
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
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



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.*



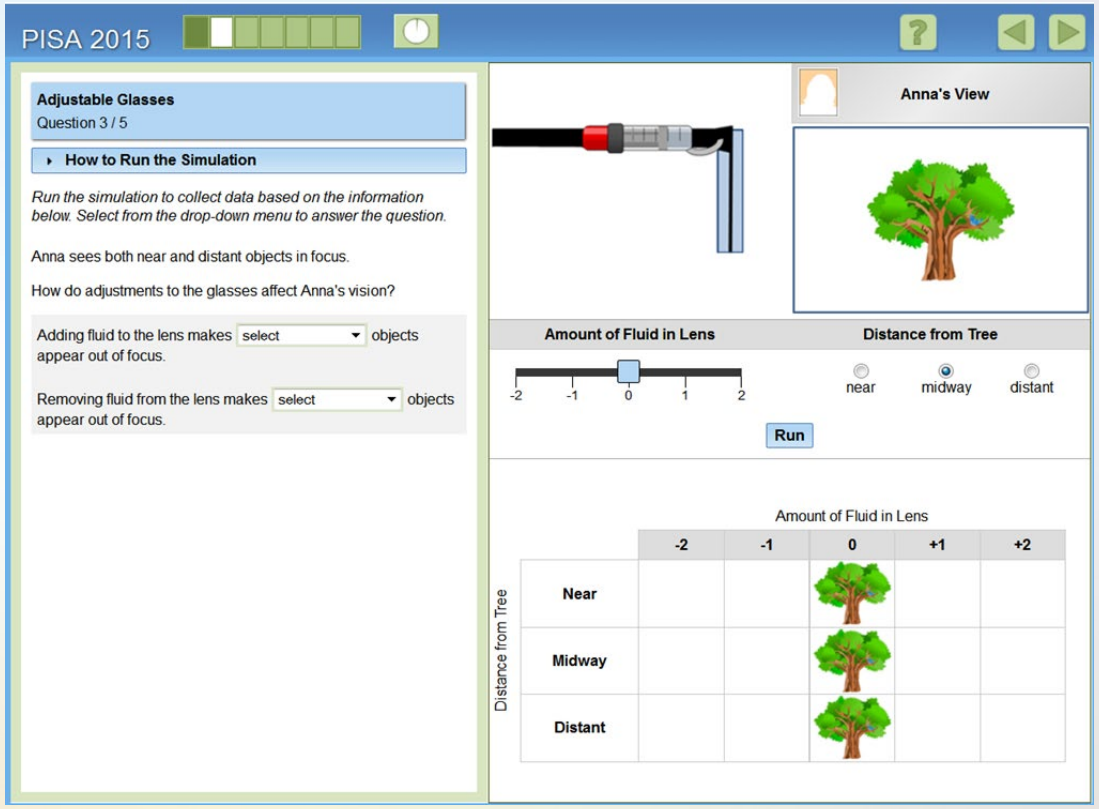
Cited from *He, Q. & Gonzalez, E. (2019). Opportunity versus Challenge: Exploring usage of log-file and process data in international large scale assessments, workshop held by Educational Research Center of Ireland and ETS, Dublin, Ireland.*

What documents do we need in the framework?

Log files	Codebook	Item files	Respondent	Features	Data files
• Raw log files • Structuralized datasets (timestamped event logs)	• Log-item codebook (by multilevel) • (Dynamic) codebook	• Item description • Scoring rules • Interface • Item calibrations and parameters • Item design / optimal problem-solving strategies • Item sequence and position	• Response data (credits) • Sample description • Background information	• Aggregated level and item level • Sequence-based features (n-grams, action sequence, etc.) • Cognitive features • Timing feature	• 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.



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