

# LOGANShiny: An app for illustrating process data analysis from international large-scale assessments

Denise Reis Costa  
Centre for Educational Measurement  
University of Oslo, Norway  
d.r.costa@cemo.uio.no

## ABSTRACT

This paper describes a Shiny application for the R package LOGAN, LOGANShiny. This app was built to provide researchers and education stakeholders an overview of basic tools for starting their analysis of process data from international large-scale assessments. Using the log file data from one item displayed at the PISA 2012 creative problem-solving assessment, the app is divided in three modules: (a) Data Preparation, (b) Response times, and (c) Respondent's actions. In each module, the user can interact with the app by analyzing students' performance on the item or comparing specific groups of students (e.g., gender or cross-country analyses). The exploration of such tools can not only illustrate the potential and limitation of process data analysis from these assessments but can also advance one's understanding of how students from 44 countries and economies interact with a problem-solving item on an international survey.

## Keywords

Computer-based assessment; Log data; Digital items; R package.

## 1. INTRODUCTION

International large-scale assessments have received widespread attention by measuring key cognitive skills and gathering information and data on how individuals use their knowledge in different contexts. For example, since 2012, two important assessments conducted by the Organisation for Economic Co-operation and Development (OECD), the Programme for International Student Assessment (PISA) and the Programme for International Assessment of Adult Competencies (PIAAC), not only started the administration of computer-based formats for a large number of participating countries but also made a number of items with respondent's log file information publicly available. These log data contain a record of the interactions between the respondents and the computer testing application during the assessment.

Process data from these kinds of data (e.g., response times and respondent's actions) are of potential relevance to researchers and can provide a better understanding of a range of issues related to test-taking behavior (e.g., engagement [3], navigation behavior [5]). Despite these potentialities, research on this field is still not well developed due to the challenges and obstacles associated with the management of such data [4].

To overcome this difficulty, an open-source R package was developed: LOG file ANalysis in international large-scale assessments (LOGAN [10]). This package is intended to present a set of user-facing functions, and the user does not need to be knowledgeable of the details of the underlying code or extensively work on the data management to conduct specific analysis of the log files from these assessments.

Figure 1 presents the LOGAN package architecture with examples of functions related to each step of the analysis of process data. First, users import their log file data into R. For PISA 2012 log files, the package has a specific function to import the semi-processed SPSS files that are freely available at the OECD website (<https://www.oecd.org/pisa/pisaproducts/database-cbapisa2012.htm>). In the future, there is also an intention to support the data management of raw log file data (e.g., xml files). After data import, one can use LOGAN functions to manage and clean the data, extract information such as the total time on the tasks or specific respondent's strategy.

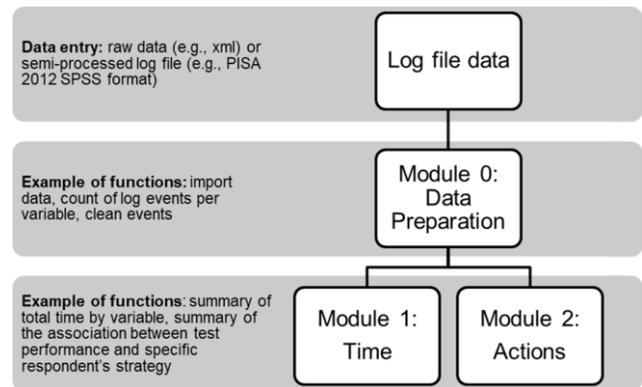


Figure 1. LOGAN package (version 1.0.0) architecture.

To demonstrate the functionalities of the LOGAN package for use by researchers and education stakeholders interested in process data analysis, a web-based application using the Shiny app [1] was created, the LOGANShiny app. Hosted at the <https://loganpackage.shinyapps.io/shiny/> page, this interactive platform brings to the users examples of analysis for one released PISA 2012 creative problem-solving item, the Climate Control (CP025Q01).

To answer this item, students were first presented a stimulus (Figure 2) where they needed to manipulate input variables (top, central, and bottom controls/sliders) to understand how an air conditioner changes the temperature and humidity of a room. Then, students had to draw arrows on a diagram that represent the relationship between the three controls and the two outputs (temperature and humidity). Full credit was given to students who correctly completed this diagram (i.e., top control impacts temperature and central and bottom controls on humidity).

The available log file data from this item captured the student's time on the task, and their exploration on applying and resetting the input variables using the sliders, the associated temperature and humidity values, and the state of the diagram at each exploration. There was no restriction on the number of times a student could manipulate these features, and they did not change

by themselves without the student's interaction. Example of studies using process data from this item are: [2], [4], [6], [7], [9], and [12].

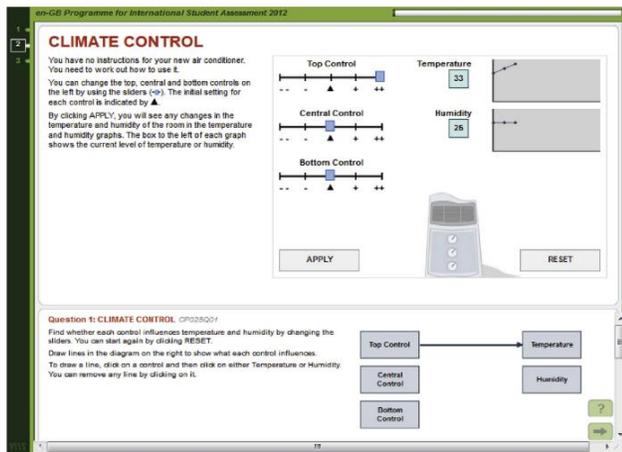


Figure 2. Stimulus information from the problem-solving Climate Control unit and the CP025Q01 item (Reprinted from [4] with permission from Elsevier).

The following sections of this paper intend to showcase the features of the LOGANShiny app regarding data management and statistical analysis from process data for this item.

## 2. DATA PREPARATION (MODULE 0)

In tab “Module 0” from LOGANShiny app, a user will be presented with the particularities of the Climate Control item and its related log-file data. An interactive table displays all the 13 variables and 951,481 entries existing in the data. It represents how students from 44 countries and economies interacted with this item. The description of each variable and the three log events of one student from the United Arab Emirates (code = “ARE”) is illustrated in Table 1.

Table 1. Log data variables description with three log events from student ID = “04852” from the “ARE” country.

Variable	Description	Log events (example)		
		1	2	3
cnt	Country code	ARE	ARE	ARE
schoolid	School ID	0000189	0000189	0000189
StIDStd	Student ID	04852	04852	04852
event	Event status	START_ITEM	ACER_EVENT	ACER_EVENT
time	Event time (in seconds)	1288.1	1291.9	1338.4
event_number	Even sequence number	1	2	3
event_type	Event type	NULL	reset	apply
top_setting	Slider position: top	NULL	0	1
central_setting	Slider position: central	NULL	0	1
bottom_setting	Slider position: bottom	NULL	0	1
temp_value	Temperature value	NULL	25	27
humid_value	Humidity value	NULL	25	28
diag_state	Diagram status	NULL	NULL	NULL

The first log event for the student ID = “04852” indicates when the student was exposed to the item for the first time (event status= “START\_ITEM”). For this case, the registered time was 1288.1 seconds since the beginning of the assessment, and no interactions with the item features were recorded (e.g., top\_setting= “NULL”).

Less than four seconds after starting the item, this student clicked on the “RESET” button. In this scenario, all the input variables were set as 0 (indicated by a triangle in Figure 2), the output indicated “25”, and no arrows were drawn on the diagram as the default.

About 46 seconds after resetting the task, the same student moved all sliders in one position to the right (i.e., top\_setting= “1”, central\_setting= “1”, and bottom\_setting= “1”) and clicked on “APPLY”. In this case, the temperature value automatically changed to “27” and the humidity to “28”. Again, the status of the diagram was still in its initial setting.

When the student interacted with the diagram, the information displayed in the “diag\_state” variable was represented as a binary number (e.g., “000001”) with each digit associated with one input and one output variable (e.g., top control and temperature).

After looking at the information that one can extract from the available log file data, analytical tools are presented at the LOGANShiny app. For example, a summary of the total number of event actions (including “START\_ITEM” and “END\_ITEM”) can be performed in the app. Figure 3 illustrates the log events from 1,015 students from Bulgaria (code=“BGR”). The same analysis can also be done for data from other countries.

### Summary of students' event actions by:

BGR

Summary of number of events by country - Individual level						
cnt	Total N	Min	Average	SD	Median	Max
BGR	1015	1	21.89	20.71	16	193

Figure 3. Interactive summary table of the number of event actions (including “START\_ITEM” and “END\_ITEM”) for students from the “BGR” country.

From the provided summary statistics, one can verify issues related to the OECD log data. For example, a student from this country has only one entry in the log file data (i.e., one event action). Since it is expected to have at least two events for this dataset (“START\_ITEM” and “END\_ITEM”), the app emphasizes the importance of a closer look at the data and acknowledges researcher’s freedom to review/filter/delete such inconsistencies.

After data management, two analytical tools are provided in the subsequent tabs: Time (Module 1) and Actions - Cognitive related (Module 2).

## 3. RESPONSE TIMES (MODULE 1)

On this tab, the amount of time students spent on the Climate control item is analyzed. First, a user should decide if the analysis of the total time will be conducted by item performance (CP025Q01=0: incorrect answer; CP025Q01=1: correct answer) or gender (ST04Q01=1: female; ST04Q01=2: male). Later, the user can choose if the analyses will consider all countries or select a specific country, as illustrated in Figure 4.

#### Analysis of time on task by:

- Item Performance  
 Gender

#### Select the level of the analysis:

Country: ALL ▼

Figure 4. Type of analytical tools presented at “Module 1”.

After these choices, two types of descriptive statistics are provided: a summary table and a density plot.

### 3.1 Summary table of response times

Figure 5 displays the information that one can gather from the LOGAN package for the analysis of the overall time and by item performance. A total of 30,345 students from all PISA 2012 participating countries and economies was analyzed for the Climate control item. The maximum amount of time spent on this item was 26 minutes from a student who got an incorrect answer. For the group of students who got a correct answer, the maximum was 16 minutes.

In general, students spent an average of 2 minutes on the task. However, this estimate is not precise since negative response times were observed in this sample (i.e., the minimum amount of time for those who got an incorrect answer was equal to -0.43). Although one could remove such cases from the dataset as they lower the average values, it is displayed in the LOGANShiny to reveal another inconsistency in this log-file dataset.

Even though this type of discrepancy could possibly be detected at the data preparation stage of analysis, it was left to the “Module 1: Time” tab for the LOGANShiny to highlight again the importance of further inspection of process data and proper data manipulation of the files from large-scale assessments.

Summary of Time on Task by CP025Q01 - Individual level			
Statistics	ToT Total	CP025Q01=0	CP025Q01=1
Total N	30345	15269	15076
Min	-0.43	-0.43	0.16
1st.Qu	1.35	1.11	1.54
Median	1.91	1.76	2
Mean	2.09	1.98	2.2
SD	1.17	1.31	1
3st.Qu	2.6	2.56	2.63
Max	26.67	26.67	16.84

Figure 5. Interactive summary table of the overall time on the task (in minutes) and by item performance for students from all participating countries and economies from PISA 2012 problem-solving assessment.

### 3.2 Time density plot

Figure 6 illustrates the distribution of time by the performance on the task obtained with the LOGANShiny app. One could also plot the distribution of time by gender.

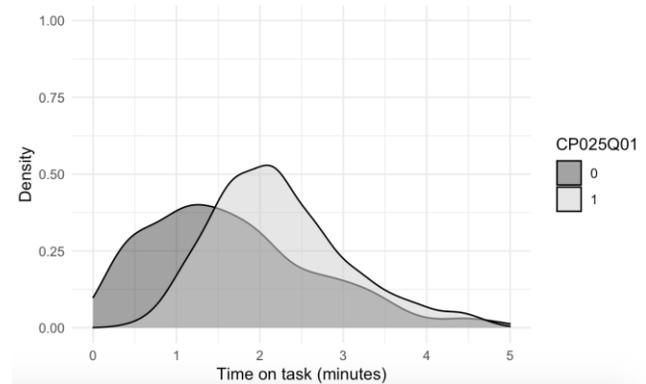


Figure 6. Distribution of the total time students from all PISA 2012 participating countries and economies spent on the item.

### 4. RESPONDENT’S ACTION (MODULE 2)

To illustrate how to explore the actions recorded in the log files, LOGANShiny describes two respondent’s action strategies discussed at [4] based on the vary-one-thing-at-a-time (VOTAT) strategy. In the case of the Climate control item, the VOTAT strategy consists of a student varying one specific variable (i.e., put the top control on “++”), while keeping all other variables constant (i.e., put the central and bottom controls on the delta symbol), and clicking on “apply”. To operationalize the VOTAT strategy, [4]’s authors suggest:

- VOTAT 1: a dichotomous variable with “1” to students who applied VOTAT for all input variables; and
- VOTAT 2: incorporated four categories for no isolated variation at all (category 0), isolated variation of one input variable (for example, only the top control), isolated variation of two input variables (for example, the top and bottom controls), and isolated variation of all three input variables (category 3).

One must note that VOTAT 2 category 3 is the same as the VOTAT 1 = “1”. Category 0, on the other hand, indicates the case where the student did not vary any slider or vary all the sliders at a time before clicking the “apply” button.

To illustrate how to derive these VOTAT variables from the log data, the third log event from Table 1 shows the case where the student selected top\_setting= “1”, central\_setting= “1”, and bottom\_setting= “1” before clicking on apply. In this scenario, both VOTAT 1 and VOTAT 2 would receive the value “0” where no isolated variation on all the controls were found.

Based on these categories, one can investigate how performance outcomes and VOTAT strategies are related by country, item level performance (CP025Q01=0: incorrect answer; CP025Q01=1: correct answer), and problem-solving overall performance (first plausible value, PVICPRO). To do this on LOGANShiny, one should select the type of VOTAT strategy they are interested in, followed by each participating country the analyses will be related to (Figure 7):

Choose a strategy

VOTAT1

Analysis by:

Country: ALL

Figure 7. Type of analytical tools presented at “Module 2”.

After these choices, two types of descriptive and correlational statistics are provided: a summary report and a frequency plot.

### 4.1 Summary report of student’s strategies and performance

On LOGANShiny, it is possible to conduct a statistical summary of students’ exploration via the “VOTAT 1” strategy and its relationship with performance. This analysis is presented as a report divided in three parts: (1) frequency table, (2) measures of association between strategy and item performance, and (3) summary of test performance (considering the first plausible value from the PISA 2012 problem-solving assessment) by VOTAT strategy. Figures 8 and 9 show an example of this report.

From Figure 8, it is possible to see that about half of the students from this sample applied the VOTAT 1 strategy at least one time. For the group of students who got a correct answer in the CP025Q01 item, the majority (12,404 out of 15,076 students) applied this strategy at least once during the item evaluation. Correlational measures (i.e., chi-square statistic and phi coefficient) are also provided to evaluate the strength of the association between these variables.

Frequency table - Individual level				
Construct	Scale	Relat. Freq.	Frequencies	Total N
Strategy: VOTAT1	0 1	52% 48%	15756 14589	30345
Performance: CP025Q01	0 1	50% 50%	15269 15076	30345

Measures of association between VOTAT1 and CP025Q01 - Individual level			
VOTAT1 / CP025Q01	0	1	Total
0	13084	2672	15756
1	2185	12404	14589
Total	15269	15076	30345

Chi-squared = 14035.15, df = 1, p-value < 0.01  
Phi coefficient = 0.68

Figure 8. First part of the interactive summary report with the analysis of student’s strategy “VOTAT 1” and item performance of all students from the PISA 2012 problem-solving assessment.

Imputation methods are used in PISA to generate plausible values to report students’ overall performance [8]. In a scale with a mean score among OECD countries of 500, five plausible values were defined for the PISA 2012 creative problem-solving assessment. In the LOGANShiny, an analysis using one plausible values is illustrated in Figure 9.

Summary of PV1CPRO by VOTAT1 - Individual level			
Statistics	Total	0	1
Total N	30345	15756	14589
Min	35.49	35.49	141.01
1st.Qu	410.47	360.96	502.18
Median	489.2	420.21	553.32
Mean	484.26	420.4	553.22
SD	107.87	92.32	76.39
3st.Qu	562.24	481.89	604.45
Max	836.58	764.34	836.58

Biserial/Polyserial correlation = 0.7714

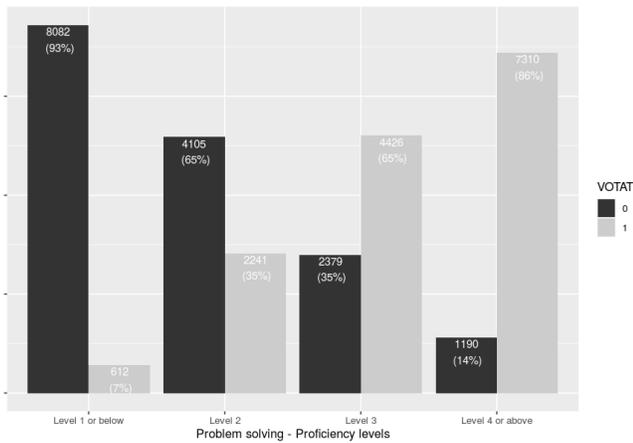
Figure 9. Second part of the interactive summary report with the analysis of student’s strategy “VOTAT 1” and overall performance (first plausible value, “PV1CPRO”) of all students from the PISA 2012 problem-solving assessment.

Based on the provided statistics, it is possible to note that students who used the VOTAT 1 strategy on the Climate Control item received, on average, more than 100 score points on the PISA 2012 creative problem-solving assessment in contrast to those who did not use this strategy.

### 4.2 Frequency Plot

In PISA, student’s scores in the assessments are also divided into proficiency scale levels to provide a substantive meaning of the overall performance. For PISA 2012 creative problem-solving assessment, seven levels of proficiency were created where level 1 (358 < PV1CPRO ≤ 423) corresponds to an elementary level of problem-solving skills and level 6 (PV1CPRO ≥ 683) the highest level. A complete description of these levels is presented in Figure V.2.2 from the OECD report [8].

In LOGANShiny, these proficiency levels are plotted in relation to the use of the VOTAT strategy. Figure 10 illustrates this relationship. Here, percentages within the categorized proficiency score are provided in parenthesis for each PISA proficiency level. Findings from this analysis indicate that students on the high level of the scale tend to use “VOTAT 1” more than those on the lower levels of the PISA 2012 creative problem-solving proficiency scale.



**Figure 10. Frequency of students by “VOTAT 1” strategy and PISA proficiency levels for students from all participating countries and economies in the PISA 2012 problem-solving assessment.**

## 5. CONCLUSION

In this paper, LOGANShiny is presented as an illustrative tool for showcasing the functionalities of the LOGAN R package functions for the analysis of process data from international large-scale assessments. Interactive tables and graphical displays intended to shed light on the potentialities and limitations of the use of log-file data regarding data management and analysis of response times and student’s actions. This app can be a valuable tool to deepen researchers’ and education stakeholder’s knowledge on the item features and provide insights on students’ cognitive process. The understanding of how process data can be extracted and analyzed may not only inspire the development of new item features that could enrich one’s experience with digital environments, but also has the potential to improve the assessment’s results by, for instance, incorporating process data into the scoring procedure [11].

## 6. REFERENCES

[1] Chang, W., Cheng, J., Allaire, J., Xie, Y., & McPherson, J. (2020). shiny: Web Application Framework for R. R package version 1.4.0.2. <https://CRAN.R-project.org/package=shiny>

[2] Chen, Y., Li, X., Liu, J., & Ying, Z. (2019). Statistical Analysis of Complex Problem-Solving Process Data: An Event History Analysis Approach. *Frontiers in Psychology*, 10, 486. <https://doi.org/10.3389/FPSYG.2019.00486>

[3] Goldhammer, F., Martens, T., & Lüdtke, O. (2017). Conditioning factors of test-taking engagement in PIAAC: an exploratory IRT modelling approach considering person and item characteristics. *Large-Scale Assessments in Education*, 5:18. <https://doi.org/10.1186/s40536-017-0051-9>

[4] Greiff, S., Wüstenberg, S., & 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. *Computers and Education*, 91, 92–105. <https://doi.org/10.1016/j.compedu.2015.10.018>

[5] Hahnel, C., Goldhammer, F., Naumann, J., & Kröhne, U. (2016). Effects of linear reading, basic computer skills, evaluating online information, and navigation on reading digital text. *Computers in Human Behavior*, 55, 486–500. <https://doi.org/10.1016/j.chb.2015.09.042>

[6] Han, Z., He, Q., & von Davier, M. (2019). Predictive Feature Generation and Selection from Process Data in PISA Simulation-Based Environment: An Implementation of Tree-based Ensemble Methods. *Frontiers in Psychology*, 10, 2461. <https://doi.org/10.3389/fpsyg.2019.02461>

[7] He, Q., & von Davier, M. (2016). Analyzing Process Data from Problem-Solving Items with N-Grams. In *Handbook of Research on Technology Tools for Real-World Skill Development* (pp. 750–777). <https://doi.org/10.4018/978-1-4666-9441-5.ch029>

[8] OECD. (2014). PISA 2012 Results: Creative Problem Solving (Volume V): Vol. V. OECD Publishing. <https://doi.org/https://doi.org/10.1787/9789264208070-en>

[9] Pejic, A., & Molcer, P. S. (2016). Exploring data mining possibilities on computer based problem solving data. *SISY 2016 - IEEE 14th International Symposium on Intelligent Systems and Informatics, Proceedings*, 171–176. <https://doi.org/10.1109/SISY.2016.7601491>

[10] Reis Costa, D., & Leoncio, W. (2019). LOGAN: An R package for log file analysis in international large-scale assessments. *R Package*. <https://cran.r-project.org/web/packages/LOGAN/index.html>

[11] Reis Costa, D., Bolsinova, M., Tilmstra, J., & Andersson, B. (2021). Improving the Precision of Ability Estimates Using Time-On-Task Variables: Insights From the PISA 2012 Computer-Based Assessment of Mathematics. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.579128>

[12] Xu, H., Fang, G., Chen, Y., Liu, J., & Ying, Z. (2018). Latent Class Analysis of Recurrent Events in Problem-Solving Items. *Applied Psychological Measurement*, 42(6), 476–498. <https://doi.org/10.1177/0146621617748325>