Response Process Data in Educational and Psychological Assessment: A Scoping Review of Empirical Studies

Guanyu Chen
The University of British Columbia
cguanyu@student.ubc.ca

Yan Liu
Carleton University
YanLiu5@cunet.carleton.ca

ABSTRACT

With the advance of computerized assessments, response process data (RPD) become available. RPD has been increasingly gaining popularity because it can help to understand and study the cognitive processes of test takers. We aim to conduct a scoping review to provide a comprehensive overview of the common practice and major findings with a focus on the theoretical framework and analytical methods applied in RPD studies. This review can help researchers understand the advantages and challenges of using RPD in both educational and psychological fields. Our findings provide guidance to researchers who are interested in RPD applications.

Keywords

scoping review, response process data, log file data, assessment,

1. INTRODUCTION

With the recent development of computer technology, response process data (RPD) are widely collected in computerized assessments [1]. RPD reflect the thinking processes, strategies, and behaviors of test takers when they read, interpret, and formulate solutions to assessment tasks [2]. RPD can document test-taking behaviors that may not be observed directly from test scores, which can show response patterns and thinking processes and may possibly provide learners and other stakeholders with more meaningful feedback [3]. Given the importance of RPD, the purpose of this scoping review is to examine the extent, range, and characteristics of RPD, to summarize analytical methods used as well as the findings obtained from application studies, and to identify gaps in the literature [4].

2. RESPONSE PROCESS DATA

RPD can be traced back to the log files, which record events that occur in a computer system [5]. RPD is one type of log-file data, also known as (response-related) paradata in survey research [6], recording the interactions between the test takers and the computer [7]. In computer-based assessment contexts, both the test takers' actions to the stimulus materials and the ordered sequence (i.e., the timestamps) of these actions are stored in RPD [2], [3], [8].

RPD are usually stored in a structured format, such as XML and JSON, and RPD need to be parsed and converted into a tabular

sequence number. The fourth column provided detailed information (i.e., properties) about the event.

Table 1. Process Data from PISA 2012 Problem Solving

event	time	event_number	event_value
START_ITEM	0.10	1	NULL
ACER_EVENT	43.40	2	'00000000000001000000 0000
click	43.40	3	hit_nowhereSakharov
ACER_EVENT	44.90	4	'00000000000000000000 0000

3. STUDY PURPOSES
3.1 Scoping Review

Although RPD is an emerging topic and there are a number of empirical studies that have been conducted, there is no review that has been carried out to offer insights into the current applications related to RPD according to our best knowledge. A scoping review maps the key concepts behind a research topic and different sources of evidence, and the scoping review can be conducted as a stand-alone study, especially for a complex and emerging topic [10]. Conducting a scoping review will contribute to an overall understanding of the current application of RPD across different research areas in educational and psychological assessment.

Specifically, we will undertake the scoping reviews for examining the extent and characteristics of research with RPD. It is important to gain insights into how RPD are being applied and analyzed as a gold mine in educational and psychological assessment. By summarizing the current research, theoretical and analytical frameworks for RPD will be identified and examined for providing a broader overview of these indicators, methods, and findings. Finally, this scoping review could also be used to guide further research and practice.

3.2 Review Objects

This scoping review aims to systematically investigate how RPD are being used in educational and psychological assessment for answering the following questions:

(1) What is the theoretical framework that supports the analysis of the RPD?

(2) How to extract and generate suitable indicators from the raw RPD? And what kind of indicators have been used in the current practice?

(3) What analytical methods have been used for RPD?

(4) Based on the indicators and corresponding methods used in the existing studies, what inferences have been made? More

specifically, what are the study purpose and corresponding findings?

4. METHODS

4.1 Study Design

We adapted Arksey and O'Malley's framework to organize this scoping review. Additionally, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) provided the checklist for essential reporting items [4], [11], which were also used to design our scoping review.

4.2 Search Strategy

The search query for the Web of Science was provided here with the consideration of the Peer Review of Electronic Search Strategies (PRESS) checklist [12]:

TS = ((paradata OR "process data" OR "log data" OR "log-file data" OR "logfile data" OR "mouse click" OR keystroke OR keypress) AND (survey OR questionnaire OR "test batter*" OR assessment OR PISA OR PIAAC OR NAEP OR TIMSS OR PIRLS)) AND PY = (2000-2022)

The query returned 1904 records in Web of Science and around 5000 records from all databases, including ERIC, Education Source, PsycInfo, ProQuest Dissertations & Theses Global, Web of Science, and Scopus.

4.3 Inclusion and Exclusion Criteria

As a scoping review, we include all types of empirical research, including gray literature from ProQuest Dissertations & Theses Global. Most of the studies use RPD as a secondary data analysis. Thus, we expect very few experimental studies, and most of the studies will be observational studies. As mentioned in Research Question Section, we focused on the theoretical and analytical frameworks of RPD in practice. Hence, we excluded methodological studies which focused on the simulation or algorithm. Review studies will be considered, and the empirical studies included in review studies will be retrieved and reviewed. However, we can only include full-text and English articles for conducting the full-text review according to the background of reviewers. Finally, this scoping review includes all human populations in any context as long as their interactions with computers were recorded.

4.4 Study Selection

Study selection is an iterative, rather than linear, stage involving a process of searching the literature, refining the search strategy, and reviewing articles for study inclusion and exclusion criteria [13]. At least two independent reviewers were asked to perform the study selection for the title and abstract screening and full-text screening. Another content expert was invited to solve the disagreement between the reviewers. Some pilot tests were recommended before the formal selection for refining this study selection process [11]. We will choose a sample of 50 articles, review these articles with eligibility criteria, discuss the discrepancies, and modify the search query and eligibility criteria.

We will use a flowchart of the review process from PRISMA-ScR to describe the whole scoping review process, including the databases, duplications, screening, full-text retrieval, and additional search from reference lists and relevant organizations. Covidence will be used for data management and screening.

5. EXPECTED RESULTS

5.1 Data Extraction

Google Forms will be used for developing the data charting form to collect the information for answering research questions. A series of key information will be recorded, such as:

(1) Citation information: author(s), publication year

(2) Indicators: generation, definition, type, theoretical framework

(3) Methods: name, category

(4) Inferential framework: aim of the study, findings

Note that additional information will be included during the review, and the chart form will be continually updated. After the review team discusses and trials the chart form and the chart form, two independent reviewers will extract the information to ensure the accuracy of data extraction.

5.2 Data Synthesis

To clarify our results, we will break our data synthesis into three steps [13]. First, we need to conduct the data analysis. The frequency counts of indicators, methods, and findings are used for depicting the extent, range, and characteristics of the studies included in the scoping review [10], [14]. Moreover, to provide in-depth analyses, descriptive qualitative data analysis, such as thematic analysis with human coding [15], will be used [11]. Thematic analysis can summarize the data into a particular category (i.e., classifying the statistical methods into descriptive statistics or inferential statistics). Then, according to the research questions, we will report the results and produce the findings. A small table includes the characteristics of all the studies under a specific topic, (i.e., indicators, methods, and inferential frameworks in this review). Finally, the implications of our results will be considered with the overall research purpose and the specific research question and extended to the broader context for future research, policy, and practice [13].

6. DISCUSSION

RPD is an emerging and developing research topic in the fields of psychology and education. With the wide use of computer-based assessment, RPD becomes more and more available. However, the significantly increased volume, velocity, and variety of RPD raise new challenges for researchers to handle, analyze, and interpret them in order to materialize the value [1]. As there is a lack of scoping review to provide a comprehensive overview of the current theoretical and analytical frameworks to guide future research and practice. Even though a variety of analytic methods were used for different indicators, this scoping review will provide a systematic summary of common indicators, methods, and findings.

7. REFERENCES

[1] von Davier, A. A., Mislevy, R. J., & Hao, J. (Eds.). (2021). *Computational Psychometrics: New Methodologies for a New Generation of Digital Learning and Assessment:* With Examples in R and Python. Springer International Publishing. https://doi.org/10.1007/978-3-030-74394-9

[2] Ercikan, K., & Pellegrino, J. W. (2017). *Validation of score meaning for the next generation of assessments: The use of response processes.* Taylor & Francis.

[3] Jiao, H., He, Q., & Veldkamp, B. P. (2021). Editorial: Process Data in Educational and Psychological Measurement. *Frontiers in Psychology*, 12.

https://www.frontiersin.org/articles/10.3389/fpsyg.2021.7933
99

[4] Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine*, 169(7), 467–473. https://doi.org/10.7326/M18-0850

[5] Andrews, J. H. (1998). Testing using log file analysis: Tools, methods, and issues. *Proceedings 13th IEEE International Conference on Automated Software Engineering (Cat. No.98EX239)*, 157–166. https://doi.org/10.1109/ASE.1998.732614

[6] Kroehne, U., & Goldhammer, F. (2018). How to conceptualize, represent, and analyze log data from technology-based assessments? A generic framework and an application to questionnaire items. *Behaviormetrika*, 45(2), 527–563. https://doi.org/10.1007/s41237-018-0063-y

[7] OECD. (2019). *Beyond Proficiency: Using Log Files to Understand Respondent Behaviour in the Survey of Adult Skills*. OECD. https://doi.org/10.1787/0b1414ed-en

[8] OECD. (2015). *Students, Computers and Learning: Making the Connection*. OECD. https://doi.org/10.1787/9789264239555-en

[9] Hao, J., & Mislevy, R. J. (2021). A Data Science Perspective on Computational Psychometrics. In A. A. von Davier, R. J. Mislevy, & J. Hao (Eds.), *Computational Psychometrics: New Methodologies for a New Generation of Digital Learning and Assessment* (pp. 133–158). Springer.

[10] Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19–32. https://doi.org/10.1080/1364557032000119616

[11] Peters, M. D. J., Marnie, C., Tricco, A. C., Pollock, D., Munn, Z., Alexander, L., McInerney, P., Godfrey, C. M., & Khalil, H. (2020). Updated methodological guidance for the conduct of scoping reviews. *JBI Evidence Synthesis*, 18(10), 2119. https://doi.org/10.11124/JBIES-20-00167

[12] McGowan, J., Sampson, M., & Lefebvre, C. (2010). An Evidence Based Checklist for the Peer Review of Electronic Search Strategies (PRESS EBC). *Evidence Based Library and Information Practice*, 5(1), Article 1. https://doi.org/10.18438/B8SG8R

[13] Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, 5(1), 69. https://doi.org/10.1186/1748-5908-5-69

[14] Peters, M. D. J., Godfrey, C., McInerney, P., Munn, Z., Tricco, A. C., & Khalil, H. (2020). Chapter 11: Scoping reviews (2020 version). In A. E. & M. Z. (Eds.), *JBI Manual for Evidence Synthesis (Vol. 2020)*. JBI. https://doi.org/10.46658/JBIMES-20-12

[15] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.