Researching education, improving learning

Amsterdam, April 19, 2021

CALL FOR PROPOSALS

Call no. IEA 13/04-2021

IEA Research for Education Series: Setting a Foundation for the Interpretation and Use of Process Data in International Assessment

1. Introduction

The International Association for the Evaluation of Educational Achievement (IEA) invites proposals for creating a report based issues fundamental to IEA studies. The general theme for this report is around developing a general framework for using process or log file data stemming from a computer-based platform. The deliverable for this project will be an 80- to 150-page book, to be published by Springer as part of IEA's Research for Education Series (see https://link.springer.com/bookseries/14293). The book will include, in addition to the main text, tables, graphs, cited references, and relevant appendix materials.

2. Study background and objectives

With the advent of technology, IEA studies are undergoing a steady transition from a paper-and-pencil to a computer-based platform. This technological shift heralds a new and massive class of *process* data that can be extracted from computer-generated log files. Although definitions of process data differ (Ercikan & Pellegrino, 2017; Hubley et al., 2017), they generally refer to *the way a student* arrives at an answer or response. These data include the amount of time spent on a page, the number of clicks or keystrokes, page returns, the time and date an application was accessed, and so on. Innovative uses of process data include forensic methods of detecting data fabrication (Yamamoto & Lennon, 2018) and test taking pattern identification (Teig et al., 2020), among others. Further, the sheer volume of data associated with log files necessitates advanced methods for data processing (Goldhammer et al., 2020) and analysis (machine learning or topic modeling, as two examples). In spite of the growing ubiquity of process data and an increased recognition by the research community that these data are ripe for harvesting, essential questions around what *can* and *should* be done with these data exist. Taking a higher-level view on the matter, it is important to carefully consider *what log file data should be collected, processed, and made available to researchers.*

One key area around which these questions should be answered is the intersection of psychometrics, which historically focused on item responses or *product data* (Bergner & von Davier, 2019; Levy, 2020) and process data. Importantly, a number of methods exist for integrating *product* measurement with *process* measurement, including techniques for jointly modeling response accuracy and response time (van der Linden, 2007; van der Linden et al., 2010), among others. To that end, a small but growing body of literature is emerging in this area, with much of the work emphasizing Mislevy's (Mislevy et al., 2003) Evidence Centered Design framework (Goldhammer & Zehner, 2017; Levy, 2020). Further, Zumbo and Hubley recently published an entire edited volume on issues around using response processes as validity evidence; however, just one chapter dealt with the international assessment context directly (Chen & Zumbo, 2017). To that end, we seek a volume that specifically addresses the context of IEA studies and whether and how process data should be collected and whether and how these potentially rich data can be used. In particular, IEA seeks a volume that deals with the conceptual, theoretical, and possibly ethical issues around collecting and using process data in international assessment and the kinds of validity evidence that must be brought to bear to draw valid interpretations. The international assessment context is unique in that dozens of highly heterogeneous educational systems that differ in culture, language, geography, and economic development participate. Further, stakes at the student level are low; however, student motivation likely varies for a host of reasons. Taken together, these settings raise issues around the comparability of processes like response time or number of actions and the kinds of evidence that would be needed to base inferences on process data across educational systems. As one issue we highlight that although students are generally aware that their product data are being collected, no efforts are made to inform them that their process data are also collected, even if collaterally. Ideally, the final product should provide an authoritative source for researchers interested in using process data to draw inferences about students, their teachers, or their schools. It might also inform future operational work and, ideally, a set of international standards analogous to the AERA/APA/NCME Standards for Educational and Psychological Testing (American Educational Research Association et al., 2014), which governs testing in the U.S.

3. Possible topics

Given the IEA's aim to offer a guiding volume on the collection and use of process data, topics that could be included in this volume are:

- An up-to-date literature review of process data as validity evidence;
- A grounded argument for appropriate and inappropriate uses of process data, including in psychometric models and in other research settings;
- Theorizing around necessary evidence for making claims with process data;
- Ethical considerations around using process data, especially for making inferences about what students know and can do;
- Dominant theories around process data from fields like cognitive science, educational psychology, or learning science.

The IEA welcomes proposals that address this topic from different theoretical, conceptual, and cultural perspectives. Importantly, the cross-cultural nature of international assessment should figure prominently in the proposal. Competitive proposals should clearly describe the overarching framework and provide well-defined approaches or perspectives, and a detailed description of the expected final product. If relevant, successful author teams are expected to provide exemplar syntax and data, so that readers can replicate and extend the analyses presented in the book.

4. Data

Although this commissioned volume does not identify a single data source as the focus of this volume, IEA studies will be a central focus of a successful proposal. These studies include, but are not limited to, TIMSS, PIRLS, ICCS, and ICILS.

5. General guidelines for proposal submission[1]

Proposals must be submitted in English.

Please ensure the proposal demonstrates familiarity with the proposed research by including a sound literature review. Ensure that the contribution of the proposed thematic report to this literature is explicit, especially in terms of its potential to expand the current state of research and knowledge.

When preparing a proposal, please clearly specify the research relevance and the policy relevance of the research questions and methods selected. This specification needs to expand on and augment or complement the outline ideas set out in this call for proposals.

The proposal must furthermore describe the general analytical framework that will guide not only analyses of the IEA data but also interpretation of the results of those analyses. The description of the framework must be such that it clearly shows how the proposed analysis will address the policy-relevant research questions. The description should therefore identify:

(i) which IEA data (study, questionnaire items, indices, or constructs from questionnaires) you intend to use,

(ii) any non-IEA data sources that will be included, and

(iii) any additional data collection that is deemed necessary (such as system-level characteristics).

Please make sure that a clear and complete description of the types of quantitative or qualitative analyses to be used is included. The degree to which the methods are suited to answer the research question is an important evaluation criteria for all proposals.

[1] Please note that the nature of this call might make some of these requirements irrelevant.

In addition, the proposal must include a detailed timeline for all analyses and report-writing activities, and a well-considered budget proposal to complete the project.

When developing timelines, assume a start date of July 1, 2021 and an end date of November 1, 2022; the final manuscript of the book must be ready for print production by May 1, 2022. Although there may be a certain degree of flexibility in the timeline, it must make provision for (i) submission of a complete draft report by February 1, 2021 for review by IEA's Publications and Editorial Committee (PEC), and (ii) time for subsequent revision and language editing of the report. The corresponding author must be available for consultation with Springer Publishers during the print production period.

Budgets must include the expected number of work days needed to complete each activity related to the project and a total budget in euros or US dollars. The total budget should not exceed 25,000 euros.

The call is open to all researchers, excluding teams from IEA International Study Centers. For the latter, direct assignments are possible.

The proposal should be no more than 10 pages in length.[2] Please also provide a short (500-word maximum) biographical note on each person in the team tendering for the project. Please highlight the relevance of each person's experience to the proposed activities.

IEA will review all proposals according to their methodological quality, research and policy relevance, and budget. All tenderers will be informed of the outcome of these deliberations by late June 2021.

Proposals should be submitted by email to secretariat@iea.nl

The deadline for proposals is 1:00 p.m., Monday, May 31, 2021.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing 2014*. AERA.

Bergner, Y., & von Davier, A. A. (2019). Process data in NAEP: Past, present, and future. *Journal of Educational and Behavioral Statistics*, *44*(6), 706–732. https://doi.org/10.3102/1076998618784700

Chen, M. Y., & Zumbo, B. D. (2017). Ecological Framework of Item Responding as Validity Evidence: An Application of Multilevel DIF Modeling Using PISA Data. In B. D. Zumbo & A. M. Hubley (Eds.), *Understanding and Investigating Response Processes in Validation Research* (pp. 53–68). Springer International Publishing. https://doi.org/10.1007/978-3-319-56129-5_4

[2] Times New Roman, Arial or similar, 12 point type, double spaced.

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

Goldhammer, F., Hahnel, C., & Kroene, U. (2020). Analysing log file data from PIAAC. In D. Maehler & B. Rammstedt (Eds.), *Large-scale cognitive assessment: Analyzing PIAAC data*. Springer Nature. https://library.oapen.org/bitstream/handle/20.500.12657/41286/2020_Book_Large-ScaleCognitiveAssessment.pdf?sequence=1#page=243

Goldhammer, F., & Zehner, F. (2017). What to make of and how to interpret process data. *Measurement: Interdisciplinary Research and Perspectives*, *15*(3–4), 128–132. https://doi.org/10.1080/15366367.2017.1411651

Hubley, A. M., Wu, A. D., Liu, Y., & Zumbo, B. D. (2017). Putting flesh on the psychometric bone: Making sense of IRT parameters in non-cognitive measures by investigating the social-cognitive aspects of the items. In B. D. Zumbo & A. M. Hubley (Eds.), *Understanding and Investigating Response Processes in Validation Research* (pp. 69–91). Springer International Publishing. https://doi.org/10.1007/978-3-319-56129-5_5

Levy, R. (2020). Implications of considering response process data for greater and lesser psychometrics. *Educational Assessment*, *25*(3), 218–235. https://doi.org/10.1080/10627197.2020.1804352

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). A brief introduction to evidence-centered design. *ETS Research Report Series*, *2003*(1), i–29. https://doi.org/10.1002/j.2333-8504.2003.tb01908.x

Teig, N., Scherer, R., & Kjærnsli, M. (2020). Identifying patterns of students' performance on simulated inquiry tasks using PISA 2015 log-file data. *Journal of Research in Science Teaching*, *57*(9), 1400–1429. https://doi.org/10.1002/tea.21657

van der Linden, W. J. (2007). A hierarchical framework for modeling speed and accuracy on test items. *Psychometrika*, *72*(3), 287. https://doi.org/10.1007/s11336-006-1478-z

van der Linden, W. J., Klein Entink, R. H., & Fox, J.-P. (2010). IRT parameter estimation with response times as collateral information. *Applied Psychological Measurement*, *34*(5), 327–347. https://doi.org/10.1177/0146621609349800

Yamamoto, K., & Lennon, M. L. (2018). Understanding and detecting data fabrication in large-scale assessments. *Quality Assurance in Education*, *26*(2), 196–212. https://doi.org/10.1108/QAE-07-2017-0038