

Special Feature

Advanced Technologies in Educational Assessment

Ronny Scherer

University of Oslo, Norway

Marie Wiberg

Umeå University, Sweden

Author Note

Ronny Scherer, University of Oslo, Faculty of Educational Sciences, Department of Teacher Education and School Research (ILS), Postbox 1099 Blindern, NO-0317 Oslo, Email: [ronny.scherer@cemo.uio.no](mailto:ronny.scherer@cemo.uio.no); Marie Wiberg, Department of Statistics, Umeå School of Business, Economics and Statistics (USB), Umeå university, SE-901 87 Umeå, Sweden, Email: [marie.wiberg@umu.se](mailto:marie.wiberg@umu.se)

**Please cite this paper as follows:**

Scherer, R., & Wiberg, M. (2018). Special feature: advanced technologies in educational assessment. *Behaviormetrika*, <https://doi.org/10.1007/s41237-018-0071-y>

Special Feature: Advanced Technologies in Educational Assessment

### **Introduction**

Technology is almost everywhere—even in education, an area known for its slow change (UNESCO, 2017), advanced technologies, such as artificial intelligence, machine learning, and mobile assessments, are affecting the ways we teach and learn in classrooms. Next to the challenges that arise with the integration of technology in teaching and learning contexts (e.g., Scherer, Siddiq, & Toneur, 2019; Tang & Hew, 2017), several benefits are attracting the attention of educators and policy-makers. These benefits include but are not limited to the provision of personalized learning models, the availability of data that describe the processes of knowledge and skill acquisition, and the accessibility of complex skills due to highly interactive tasks and real-world simulations (Baker & Koedinger, 2018; Quellmalz & Pellegrino, 2009). This special issue brings to attention the potential and the challenges associated with advanced technologies in educational assessment.

### **Objectives and Content**

The aim of this special issue is to summon current developments of advanced technology-based assessments, the underlying measures and psychometric models, and their applications to assess educationally relevant constructs. Key research topics, gaps, and future developments in the context of technology-based assessment are highlighted as well.

The special issue is comprised of eight papers in which the authors focus on two key issues related to technology-based assessments: (1) the approaches to representing and analyzing process data and (2) the potential of technology for the assessment of specific constructs. Table 1 presents an overview of these papers, including some contextual information. Addressing (1), Deonovic et al. (2018) explored ways to connect Bayesian Knowledge Tracing—an algorithmic approach to model learners' mastery of the knowledge

presented or tutored in a technology-based environment—with existing Item Response Theory models to improve the assessment of learning and possible progressions of learners. The authors present a unified framework to integrate these two approaches and show its performance in a simulation study. Slater and Baker (2018) extend the existing literature on Bayesian Knowledge Tracing by examining the effect of different sample sizes on its degree of error in predicting learners' mastery. From the perspective of item banking, Rights et al. (2018) bring to attention the uncertainty in scores obtained from models of item response theory. The authors propose an approach to handle this uncertainty through Bayesian model averaging. Taking an item response theory approach, Frey et al. (2018) propose to identify possible mechanisms behind missing item responses with the help of response times. The authors illustrated their approach to missingness with a dataset of almost 800 students who took a computer-based test of digital skills and compare it to common approaches. Kroehne and Goldhammer (2018) present a framework that allows researchers to conceptualize, represent, and analyze log file data that were obtained from technology-based assessments. They further suggest using finite state machines for the analysis of log file data and illustrated this approach for the response times of PISA (Programme for International Student Assessment) questionnaire data.

The second part of the special issue is comprised of three papers targeting the technological advances in the assessment of skills and engagement (2). Specifically, de Klerk et al. (2018) developed a media-based assessment of performance in a vocational domain. The simulations included in this assessment mimic real-life situations and allow researchers to trace test-takers' behavior while interacting with the tasks. The authors gathered evidence for crafting a validity and present the results from multiple perspectives in the paper. Nguyen et al. (2018) present how learning analytics can aid the understanding of study break effects on students' engagement and pass rates in virtual learning environments. In their study of more

than 120,000 undergraduate students, the authors were identified periods of procrastination with the help of log file data. Shi et al.'s (2018) paper shows the potential of a conversational intelligent tutoring system for the assessment of adults' reading comprehension. Shi et al. (2018) conclude their paper by discussing the advantages of such systems over traditional assessments of adult literacy.

### **Conclusion**

We believe that the selected papers for this special issue illustrate the enormous potential of advanced technologies for the assessment of educationally relevant variables and constructs. Specifically, the opportunities to capture not only indicators of accuracy while learners work on certain tasks but also indicators of task behavior and knowledge mastery allow researchers to draw a more detailed picture of task performance with the help of technology. At the same time, the availability of new types of data (i.e., process data) is associated with several challenges that the authors bring to attention in this special issue: First, process data require transparent and replicable frameworks for conceptualizing, representing, and analyzing them. Second, the psychometric models to describe task performance beyond the accuracy of item responses must be developed further so that process data (e.g., response times, sequences of actions, frequency of actions) can inform and improve the description of learners' proficiency (see also von Davier, 2017). Third, advanced technologies, as they allow for interactive task designs (e.g., simulations, intelligent tutoring, adaptive testing), may improve the assessment of traditional constructs (e.g., literacy and numeracy) and enable the assessment novel constructs that require complex task designs (e.g., collaborative skills, adaptive problem solving). Together with the availability of process data, however, we believe that this potential requires the careful crafting of a validity argument to facilitate the interpretation of the scores and indicators resulting from advanced technology-based assessments (see also Katz et al., 2017; Mislevy, 2016). Overall, we encourage

researchers in the field to not shy away from novel approaches to modeling and assessing complex constructs in education with the help of modern psychometrics and advanced technologies.

Table 1

*Overview of the papers presenting advances in technology-based educational assessment*

<i>Paper</i>	<i>Context</i>	<i>Advances in technology-based assessments</i>
<b>Approaches to representing and analyzing process data from advanced technologies</b>		
Deonovic, Yudelson, Bolsinova, Attali, & Maris (2018)	Connecting Bayesian Knowledge Tracing and Item Response Theory to model learner data	Psychometric modeling techniques to improve the assessment of learning and learning progressions in technology-based assessments
Slater & Baker (2018)	Errors in Bayesian Knowledge Tracing estimates due to sample size differences	Using Bayesian Knowledge Tracing to predict students' mastery in modern adaptive learning systems & the reduction of error in these predictions
Rights, Sterba, Cho, & Preacher (2018)	Using item response theory scores in item banking	Psychometric modeling techniques to handle uncertainty in item response theory scores through Bayesian model averaging
Frey, Spoden, Goldhammer, & Wenzel (2018)	Treating missing item responses in computer-based testing	Response-time based modeling approach to identify the possible mechanisms underlying missing item responses
Kroehne & Goldhammer (2018)	Framework for conceptualizing, representing, and analyzing log file data from technology-based assessments	Finite state machines to analyze log file data & the value of response times for analyzing questionnaire data from international large-scale assessments in education
<b>Advanced technologies for the assessment of specific constructs</b>		
De Klerk, Veldkamp, & Eggen (2018)	Design, development, and validation of a performance- and multimedia-based assessment of skills	Technology-based assessment with simulations of real-life problem situations & Crafting a multi-perspective validity argument for the resultant performance indicators and scores
Nguyen, Thorne, & Rienties (2018)	Effects of study breaks on educational outcomes measured by students' engagement and pass rates in virtual learning environments	Using learning analytics to measure behavioral engagement in virtual learning environments

---

Shi, Lippert, Shubeck, Fang, Chen, Pavlik Jr., Greenberg, & Graesser (2018)	Assessment of adults' reading comprehension in conversations	Using a conversational intelligent tutoring system for the assessment of adult literacy & the joint use of accuracy and response-time data
---	---	--

---

### References

- Baker, R. S., & Koedinger, K. R. (2018). Towards demonstrating the value of learning analytics for K-12 education. In D. Niemi, R. D. Pea, B. Saxberg, & R. E. Clark (Eds.), *Learning Analytics in Education* (chap. 2, pp. 49-62). Charlotte, NC: Information Age Publishing.
- De Klerk, S., Veldkamp, B. P., & Eggen, T. J. H. M. (2018). The design, development, and validation of a multimedia-based performance assessment for credentialing confined space guards. *Behaviormetrika*, <https://doi.org/10.1007/s41237-018-0064-x>
- Deonovic, B., Yudelson, M., Bolsinova, M., Attali, M., & Maris, G. (2018). Learning meets assessment: On the relation between item response theory and Bayesian knowledge tracing. *Behaviormetrika*, <https://doi.org/10.1007/s41237-018-0070-z>
- Frey, A., Spoden, C., Goldhammer, F., & Wenzel, F. (2018). Response Time-based Treatment of Omitted Responses in Computer-based Testing. *Behaviormetrika*.
- Katz, I. R., LaMar, M. M., Spain, R., Zapata-Rivera, J. D., Baird, J.-A., & Greiff, S. (2017). Validity issues and concerns for technology-based performance assessment. In R. A. Sottolare, A. C. Graesser, X. Hu, & G. Goodwin (Eds.), *Design recommendations for intelligent tutoring systems* (vol. 5, pp. 209-224). Orlando, FL: Army Research Laboratory.
- 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*, <https://doi.org/10.1007/s41237-018-0063-y>
- Mislevy, R. J. (2016). How Developments in Psychology and Technology Challenge Validity Argumentation. *Journal of Educational Measurement*, 53(3), 265-292.  
<https://doi.org/10.1111/jedm.12117>



- Nguyen, Q., Thorne, S., & Rienties, B. (2018). How do students engage with computer-based assessments: Impact of study breaks on intertemporal engagement and pass rates. *Behaviormetrika*, <https://doi.org/10.1007/s41237-018-0060-1>
- Quellmalz, E. S., & Pellegrino, J. W. (2009). Technology and Testing. *Science*, 323(5910), 75-79. <https://doi.org/10.1126/science.1168046>
- Rights, J. D., Sterba, S. K., Cho, S.-J., & Preacher, K. J. (2018). Addressing model uncertainty in item response theory person scores through model averaging. *Behaviormetrika*, <https://doi.org/10.1007/s41237-018-0052-1>
- Scherer, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, 128, 13-35. <https://doi.org/10.1016/j.compedu.2018.09.009>
- Shi, G., Lippert, A. M., Shubeck, K., Fang, Y., Chen, S., Pavlik Jr., P., Greenberg, D., & Graesser, A. (2018). Exploring an intelligent tutoring system as a conversation-based assessment tool for reading comprehension. *Behaviormetrika*, <https://doi.org/10.1007/s41237-018-0065-9>
- Slater, S., & Baker, R. (2018). Degree of Error in Bayesian Knowledge Tracing Estimates From Differences in Sample Sizes. *Behaviormetrika*.
- Tang, Y., & Hew, K. F. (2017). Is mobile instant messaging (MIM) useful in education? Examining its technological, pedagogical, and social affordances. *Educational Research Review*, 21, 85-104. <https://doi.org/10.1016/j.edurev.2017.05.001>
- UNESCO (2017, November 11). *Education systems too slow to reform, warns the IBE*. International Bureau of Education (IBE), UNESCO, Retrieved from <http://www.ibe.unesco.org/en/news/education-systems-too-slow-reform-warns-ibe>, accessed 29 October 2018.

Von Davier, A. A. (2017). Computational psychometrics in support of collaborative educational assessments. *Journal of Educational Measurement*, 54(1), 3-11.  
<https://doi.org/10.1111/jedm.12129>

### **Acknowledgements**

The editors would like to thank the journal managers and the journal staff of *Behaviormetrika* for their support in all practicalities. The editors would also like to thank the authors and reviewers without whom this special issue would have never come to pass.

This research was supported by the FINNUT Young Research Talent Project Grant (NFR-254744 “ADAPT21”) awarded to Ronny Scherer by The Research Council of Norway. This research performed by Marie Wiberg was supported by the Swedish Research Council Grant 2014-578.