Towards Strengthening Links between Learning Analytics and Assessment: Challenges and Potentials of a Promising New Bond

Dragan Gašević[1], Samuel Greiff[2], & David Williamson Shaffer[3]

[1]*Centre for Learning Analytics, Faculty of Information Technology, Monash University, VIC, Australia*

[2]*Institute of Cognitive Science and Assessment, University of Luxembourg, Luxembourg*

[3]*Department of Educational Psychology, University of Wisconsin-Madison, WI, USA*

dragan.gasevic@monash.edu; samuel.greiff@uni.lu; dws@education.wisc.edu

Abstract

Learning analytics uses large amounts of data about learner interactions in digital learning environments to understand and enhance learning. Although measurement is a central dimension of learning analytics, there has thus far been little research that examines links between learning analytics and assessment. This special issue of *Computers in Human Behavior* highlights 11 studies that explore how links between learning analytics and assessment can be strengthened. The contributions of these studies can be broadly grouped into three categories: *analytics for assessment* (learning analytic approaches as forms of assessment); *analytics of assessment* (applications of learning analytics to answer questions about assessment practices); and *validity of measurement* (conceptualization of and practical approaches to assuring validity in measurement in learning analytics). The findings of these studies highlight pressing scientific and practical challenges and opportunities in the connections between learning analytics and assessment that will require interdisciplinary teams to address: task design, analysis of learning progressions, trustworthiness, and fairness – to unlock the full potential of the links between learning analytics and assessment.

1 Introduction

By analyzing digital traces of user interaction with technology, learning analytics offer many opportunities to understand and enhance learning and the environments in which learning takes place (Lang et al., 2022).

The field of learning analytics has led to research and development activities in learning, teaching, and education more broadly that have attracted the attention of policy- and decision-makers in education. For example, learning analytic researchers have examined prediction of student success (Jovanović et al., 2021), uncovering learning strategies (Matcha et al., 2020), understanding affective states (D'Mello, 2017), and determining the role of social networks in learning (Joksimović et al., 2016; Poquet & Jovanovic, 2020). The use of learning analytics has also shown its potential to enhance both student retention (Arnold & Pistilli, 2012) and quality of feedback (Lim et al., 2021; Pardo, 2018), and to inform teaching practice (Martínez-Maldonado et al., 2022). Educational institutions have developed policies for learning analytics (Tsai et al., 2018), adoption and implementation strategies (Macfadyen et al., 2014), and principles for ethics and privacy protection (Ferguson et al., 2016; Kitto & Knight, 2019).

In spite of much promise, the field of learning analytics has three critical questions to address:

1. How can learning analytics help track learning progressions and inform assessment?
2. How can reliability and validity of learning analytics be improved?
3. How can learning analytics account for issues of diversity, equity, and inclusions in its practices and models?

These questions are particularly salient in today's world. In the digital age, work increasingly relies on the use of complex skills (Greiff et al., 2014); learning and assessment are intertwined (VanLehn, 2008); and both moral and practical concerns require expanding the workforce to include — and thus account for — marginalized groups.

In educational data mining, a cognate field to learning analytics (Baker et al., 2021), researchers have used assessment to support intelligent tutoring systems. These systems are primarily focused on skill development (Corbett & Anderson, 1994; Desmarais & Baker, 2012); however, there is a dearth of research that looks at the relationship between data and methods from learning analytics and formal assessments, whether summative or formative.

Although some scholars argue that learning analytics are inherently a form of assessment in the broadest sense (Knight et al., 2013; Milligan, 2018, 2020), existing learning analytic methods do not meet all of the criteria used in psychometrics to account for the different forms of validity in assessment (Kane, 2013; Messick, 1994, 1995). We posit that the weak connections between learning analytics and educational measurement is the likely reason for some of common concerns voiced about learning analytics and its use for student assessment (Lodge & Lewis, 2012).

There are many open challenges in learning analytics that are associated with the aforementioned three questions. It is often unclear the extent to which results are generalizable and actionable (Gašević et al., 2015). The theoretical foundations and properties of the domain being measured (*structural* aspect of validity) has not been examined thoroughly (Rogers et al., 2016; Wise & Shaffer, 2015). Little attention has been paid to reliability of data used in existing studies. Moreover, there is a considerable shortage of theoretically informed measures to meet external aspects of assessment validity across a range of skills (Milligan & Griffin, 2016). Finally, little work has systematically addressed challenges that underrepresented groups present to models used for data analysis.

Positive exchanges between learning analytics and assessment can go in both directions. Learning analytics can use tools, theories, and methods from assessment to improve its validity and reliability. But learning analytics also holds potential to offer benefits to the field of assessment (Milligan, 2020). Some early attempts to connect these two bodies of work have been made, for example when Ifenthaler and Greiff (2021) explored using trace data and data analytic techniques in assessment. Learning analytics can also be used to study existing assessment practices and to test open hypotheses in assessment research. However, there has been a notable absence of research to investigate how assessment research and practice can benefit from developments in learning analytics. Finally, the literature on assessment has long recognized issues of psychometric bias when a group of learners finds it harder to complete an assessment than another group (Jones & Appelbaum, 1989). Learning analytics is built upon data that may reflect existing systemic biases in society and education institutions, and in turn can inadvertently propagate or even amplify an unfair treatment of some groups of learners (Gardner et al., 2019; Prinsloo & Slade, 2018). Bringing learning analytics and assessment together has the potential to advance concerns of fairness and bias. However, there is a shortage of research on fairness and bias in learning analytics and let alone in analytics-based assessment.

This special issue was organized to bring together a collection of papers that addresses some of these open research questions and strengthen the links between learning analytics and assessment. We aim to explore differences in both data collection and analysis, which are conducted differently in learning analytics and established assessment procedures. The papers are organized to investigate implications of these differences, draw recommendations about how they can be addressed, and thus develop better methods in learning analytics and assessment.

2 Contributions in the special issue

Table 1 summarizes the papers that are included in this special issue. They are broadly grouped into three categories: (1) analytics of assessment; (2) analytics for assessment, and (3) validity of measurement. The papers address different issues in assessment, but each uses trace data to analyze existing practices in assessment or propose and validate new forms of assessment.

---- Insert Table 1 here ----

The first group of papers reports on the findings from four studies that use learning analytic approaches to support assessment, namely *analytics for assessment*. Two of the studies use video games for learning and assessment; the other two propose novel learning analytic approaches to supporting assessment in massive open online courses (MOOCs). Peters et al. (2021) report on the findings of a study that aimed to create and validate a new approach to assessment of intelligence – pattern completion, mental rotation, and spatial construction – using the popular Minecraft^{TM} video game. The study showed that tests administered through Minecraft^{TM} had moderate reliability (as demonstrated by Rasch models) and convergent and factorial validity. Importantly for this special issue, the study found that trace data was highly predictive of performance on intelligence tests in Minecraft^{TM} and moderately predictive of performance on conventional tests. Rowe et al. (2021) also made use of a video game – Zoombinis^{TM} – to measure implicit practices of computational thinking that students follow while playing the game. The study developed a set of machine learning classifiers trained on trace data from gameplay; the classifiers produced good accuracy in automatic detection of computational thinking practices. Dowell and Poquet (2021) propose a novel analytic approach for the assessment of socio-cognitive roles learners take in during online interactions. The approach is based on a combination of two data analytic techniques – social network analysis and group communication analysis – and is empirically shown to be able to effectively characterize socio-cognitive roles that emerge in peer interactions in a MOOC. Barthakur et al. (2021) introduce an analytic approach for assessing strategies that learners follow across multiple MOOCs within a professional development program. The approach is based on a latent class analysis, which a soft clustering technique, to identify program-level strategies through the analysis of trace data about learner interactions with resources available in a MOOC platform.

The second group of papers in the special includes four papers that focus on *analytics of assessment*. These papers propose analytic approaches that are used to examine assessment practices and answer questions about properties of existing assessments. Stadler et al.

(2020) use trace data to examine whether test-taking *behavior* is an effective indicator of the tested *ability*. Trace data from students taking tests of complex problem solving suggest that in this test, behavior is a good indicator of ability. Nicolay et al. (2021) used trace data to investigate whether students are able to transfer from knowledge acquisition to knowledge application during an assessment of complex problem solving. They show that many participants were *not* able to transfer knowledge, especially for the more complex items in the assessment. Zhang et al. (2021) propose a novel analytic approach for modeling the interaction between resilience and ability in assessments that allow for multiple attempts. The proposed analytic approach found that resilience both affected performance scores (and thus questioned validity for summative assessments) *and* created opportunities for ecologically valid measures of resilience that are not based on self-reports. Finally, Misiejuk et al. (2021) propose a learning analytic approach to investigate how students react to peer assessment. They used epistemic network analysis to show that students value specificity, justification, and constructiveness in peer assessment, but kindness is less of a priority.

The third group of papers is focuses on validity in learning analytics based on trace data and its implications for assessment. Winne (2020) discusses validity in learning analytics by examining self-regulated learning. He argues that theory plays a critical role in assuring validity of learning analytics and then analyzes factors that can confound validity, such as student agency while studying and the contrast between dynamic events in learning versus static assessment measures. Shute and Rahimi (2021) analyze the validity of a stealth assessment of creativity in a physics video game. They show that the proposed stealth assessment has good external validity (i.e., it can predict external performance measures) and that estimated creativity through stealth assessment is a good predictor of in-game performance, game enjoyment, and learning of physics content. Finally, Liu et al. (2021) report on the findings of a study that validated a formative assessment model of written reflection. They use confirmatory factor analysis based on textual features extracted from two datasets using well-known linguistic frameworks.

3 Future Opportunities and Challenges

The papers included in this special issue thus offer a rich set of contributions that illustrate the potential for strengthening the links between learning analytics and assessment. The three broad categories – analytics for assessment, analytics of assessment, and validity – highlight key areas of the potential connections between these fields: raising questions and possible avenues for future research. The contributions to this special issue are important developments for both learning analytics and assessment. However, they are best viewed as exemplary work early in the process of fostering connections between the two fields. As a result, of course, these contributions do not provide a complete picture of the possible links between learning analytics and assessment, the opportunities, and the open questions that can be investigated in the future work. In the reminder of this section, we highlight some of these key opportunities and open challenges.

3.1 Properties of Assessment and Learning Analytics

Validity is a critical property of assessment and has a strong tradition in educational research and practice (Kane, 2013; Messick, 1994, 1995). Accordingly, validity has received significant consideration in the contributions to this special issue, both in the papers that explicitly deal with validity (Liu et al., 2021; Shute & Rahimi, 2021; Winne, 2020), and in other papers that addressed issues of validity in assessment *for* learning (Dowell & Poquet, 2021; Peters et al., 2021; Rowe et al., 2021) and assessment *of* learning (Zhang et al., 2021). These contributions considered different facets of validity including construct and consequential validity (Winne, 2020), external validity (Rowe et al., 2021; Shute & Rahimi, 2021; Zhang et al., 2021), factorial and convergent validity (Peters et al., 2021), and structural and convergent validity (Peters et al., 2021).

There are several key challenges to be addressed in research on validity at the intersection of learning analytics and assessment. The papers here provide valuable illustrations of how both learning analytics and assessment can benefit from the consideration of issues pertinent to validity. However, the field still needs a clear theoretical framework to guide the consideration of validity in learning analytics. Existing examinations of validity in assessment (e.g., Kane, 2013; Messick, 1994, 1995) are frequently cited in the contributions in this special issue, and, of course, they indeed offer some useful directions. However, data used in learning analytics is *not always purposefully collected* to meet criteria for validity that are expected in conventional assessment. The role of theory, as emphasized by Winne (2020) in this special issue and previously in learning analytics (Gašević et al., 2015; Wise & Shaffer, 2015), is essential for validity. Therefore, a key open challenge is to develop a theoretical framework for validity in learning analytics that recognizes the specific properties of *in situ* data that learning analytics use. At the same time, there is a significant opportunity to harness new types of data (e.g., trace data) to inform the validity of assessments, as shown by Zhang et al. (2021), and to test properties of assessment in different context, as illustrated by Nicolay et al. (2021) and Stadler et al. (2020).

There is little research generally on properties of assessment in learning analytics such as reliability, fairness, sustainability, and developmental nature. In assessment, *reliability* means that assessments produce consistent results across similar contexts (Crocker & Algina, 2009). Some of the contributions in this paper make use of well-known approaches to reliability by focusing on inter-rater reliability to make sure the results produced by machine learning algorithms are in agreement with ratings by human experts (Rowe et al., 2021). Winne (2020) takes a step further and highlights that in learning analytics, reliability is not simply a function of a good design of a learning environment used for data collection; it is equally dependent on a learner's agency[1] and level of metacognitive knowledge[2] about learning tools that are available to them in the learning environment. If learners do not know about tools that are available in a learning environment, they are not likely to use them (Gašević et al., 2017; Winne, 2006). Thus, they will not 'produce' data that are deemed necessary to make assessment inferences about their learning.

Sustainability is also a critical dimension in the connections between learning analytics and assessment. In the assessment literature, sustainability is the extent to which an assessment is easy to implement and maintain (Beck et al., 2013). Learning analytics strives towards suitability through the use of data that are collected as a by-product of learning activities (Siemens, 2013). However, this requires addressing the concerns about reliability of data not collected expressly for purposes of assessment. If those concerns are met, learning analytics could offer strong opportunities for sustainable assessment through videogames (Peters et

[1] In this context, we define agency as "the capability to exercise choice in reference to preferences" (Winne, 2006, p. 8) and that learners-agents "act with purpose" (*idem*., p. 8).
[2] Metacognition can be defined as "one's knowledge concerning one's own cognitive processes and products or anything related to them" (Flavell, 1976, p. 232), while metacognitive knowledge can be defined as "knowledge of cognition" (Clarebout et al., 2013, p. 187).

al., 2021; Rowe et al., 2021; Shute & Rahimi, 2021) and formative assessment in online environments more generally (Dowell & Poquet, 2021; Liu et al., 2021).

3.2 Instrumentation and measurement

Data used in learning analytics are not always purposefully collected for measurement and assessment. While unobtrusive data collection allows for collection of large amounts of data, digital learning environments are not always instrumented to collect necessary data about learning processes, learning products, and skills (Gašević et al., 2015). Recent studies suggested that these limitations can be addressed with improvements in instrumentation of learning environments. For example, van der Graaf et al. (2021) demonstrated how introduction of specialized tools (e.g., planner or time) can enable the collection of granular trace data about processes of self-regulated learning. The validity of such trace data can be improved with the use of other data sources, which are established in the literature such as the use of think aloud protocols as reference points for validation of trace data about self-regulated learning (Fan et al., 2022).

Novel measurement approaches are needed to make use of historic trace data in assessment. Several promising approaches have been proposed in the literature. Milligan and Griffin (2016) propose an assessment instrument based on trace data in MOOCs to measure what they refer to as the "crowd-sourced learning" capability, namely, the capability to learn in environments with large numbers of learners. In the proposed instrument, the capability is theorized to have five levels (from novice to expert). Evidence for each level is demonstrated through indicators that are derived from trace data about learner activities. The instrument was validated using the item response theory on data collected from two different MOOCs. In a similar vein, two studies in this special issue used evidence centered design (Rowe et al., 2021; Shute & Rahimi, 2021) as a systematic and well-known approach to designing assessments (Mislevy et al., 2017). Other authors in this special issue also well-established psychometric and/or statistical techniques to validate their measurement approaches that are built upon the use of trace data (Barthakur et al., 2021; Dowell & Poquet, 2021; Peters et al., 2021). To assure scalability and wide-adoption of these novel measurement approaches, future research is needed on learning design practices to create learning tasks that can be used for learning analytics-based assessment with high validity and reliability.

Metadata about task conditions in learning environments is another essential precondition for learning analytics-based approaches to assessment. However, such metadata are often not readily available in trace data extracted from open-ended learning environments. Without such metadata, for learning analytic approaches, it is difficult to make automatic inferences about the pedagogical intent behind the use of certain features of a learning environment (e.g., whether a discussion forum is used for question-answering or problem-solving) across different contexts. Therefore, future research and development of open learning environments should include instrumentation principles and mechanisms that can support effective collection of metadata about task conditions.

3.3 Learning Progression

Developmental (also known as formative) assessment and learning analytics are designed with the aim to inform pedagogical decisions and actions such as giving students feedback (Taras, 2008) or providing additional help or resources. The field of learning analytics claims its intention to enhance learning in most of widely used definitions (Lang et al., 2022). To date, dashboards have been the most common format for presenting the results of data analysis to decision makers (Bodily & Verbert, 2017). However, studies show that in many cases dashboards are not an effective means to communicate the results of data analysis (Aguilar et al., 2021; Chaturapruek et al., 2018; Lonn et al., 2015). Partly, this is due to limitations in reliability and validity of the assessments being presented and the shortage of the suitable data for assessments (Matcha et al., 2020). The field has also not developed analytic approaches that track progression of learning and identify gaps in learning that require further attention. Recent efforts using epistemic network analysis to model learning progressions (Rolim et al., 2019; Shaffer et al., 2016) and to inform feedback and pedagogical practice (Herder et al., 2018) offer promising opportunities for analytics-based developmental assessment. Likewise, data analytic learning analytics techniques based on temporal and sequential modeling (Chen et al., 2018; Saint et al., 2022) potentially offer opportunities to track learning progression and provide formative guidance to teachers or students. Finally, work by Milligan and Griffin (2016) suggests that combining established principles from psychometrics with learning analytics techniques provides another avenue for measuring progression in developing skills and abilities.

3.4 Multimodal Data and Physical Environments

Learning analytics offers approaches that can enrich assessment practices through the use of multimodal data collected from in physical learning environments. The papers included in this special issue primarily use trace data from one data modality: online click behaviors, question answering, or written text. Multimodal learning analytics is a subfield of learning analytics that recognizes (a) learning is a multimodal phenomenon, (b) learning happens across multiple physical and digital spaces, and (c) multiple data channels (e.g., eye-tracking, mouse movements, spatial location, and physiological biomarkers) need to be taken into consideration to analyze learning as a complex process (Azevedo & Gašević, 2019; Sharma & Giannakos, 2020; Worsley et al., 2021). Future research that aims to strengthen the links between learning analytics and assessment should focus on approaches that can make use of multimodal data to address questions of validity and reliability in measurement (Fan et al., 2022; Wise et al., 2021) and perform measurements in physical and hybrid (physical and digital) learning environments.

3.5 Assessment Trustworthiness

Introduction of digital technologies in assessment is often associated with questions related to trustworthiness. Debates around dishonesty in assessments have been frequent in the context of MOOCs. The on-going COVID19 pandemic has brought more contentious debates about assessment in remote and distance education (Selwyn et al., 2021). Many schools and higher education institutions opted for different online proctoring solutions to address the questions of assessment trustworthiness (Kharbat & Abu Daabes, 2021). That prompted a considerable pushback and raised questions about the impact of such practices on student autonomy and privacy (Coghlan et al., 2021). Several approaches based on data analytic methods have been proposed that aim to identify academic dishonesty in online assessment, such as using multiple accounts to copy answers to assessment items (Alexandron et al., 2017) or communication between students during assessments (Ruipérez-Valiente et al., 2021). While the use of data analytic approaches holds some promise to address issues of assessment trustworthiness, future research needs to determine situations under which the use of such approaches is educationally justified and ethically acceptable. Future research should also investigate conditions under which privacy is protected to prevent the development of surveillance culture and unwarranted data sharing with third parties (Kollom et al., 2021; Selwyn, 2020), and thus, the erosion of trust in analytics-based assessment practices in education institutions (Tsai et al., 2021). Moreover, future research and development is needed on codes of practice that will promote ethical and privacy principles.

3.6 Fairness, Equity, and Inclusion

Learning analytics researchers have access to large datasets about student learning. When this data is used to assess student learning, it is critical that such models be *fair.* That is, all participants in the assessment must have equal opportunity to succeed, and the assessment should not be systematically biased toward or against certain groups (Gipps & Stobart, 2009). However, data used in learning analytics can be and often are reflective of structural biases that may exist in society and education institutions (Carter & Egliston, 2021; Selwyn, 2020). When data analysis models are trained on such biased data (e.g., prediction of students at risk of failing a course), the use of the results of such models can perpetuate the biases and even further deepen inequality (O'Neil, 2016).

This poses a problem for the development and validation of learning analytics-based assessments, however, because learning datasets typically contain subgroups: populations of students defined by demographics (e.g., race, native language, disability, income) or other metadata (e.g., attendance) that have relatively low numerical representation in a dataset. As researchers develop and validate assessments on such data, the models—and thus any assessments based on them—may be biased toward majority groups and thus ultimately *unfair* to subgroups. In machine learning, this is known as the *subgroup fairness problem* (Chouldechova & Roth, 2018; Mehrabi et al., 2019). Despite broad attention to issues of equity in education, there has been little systematic attention paid to subgroup fairness in learning analytics, despite the fact it has the potential to reify and even augment existing biases (Gardner et al., 2019; Mayfield et al., 2019; Sha et al., 2021).

4 Conclusion

This special issue is meant to serve as a catalyst for strengthening research links between learning analytics and assessment. We were very fortunate to assemble an outstanding group of papers that were contributed by authors with different theoretical backgrounds. The high-quality contributions included in this special issue provide a good overview of the state-of-the-art on this topic. The contributions offer important insight into the complexity of the relationship between learning analytics and assessment. Advancements in understanding of this relationship can inform future work of researchers, practitioners, and policy makers to develop novel forms of assessment and increase rigor of learning analytics. We hope that this special issue and the challenges and opportunities discussed in this editorial will inspire future work on the links between learning analytics and assessment.

5 Acknowledgements

The guest editors would like to thank Prof. Paul Kirschner, Commissioning Editor of Computers and Human Behavior, for his support of this special issue, expert advice, and leadership. The guest editors are also thankful to the journal managers and the journal staff of Computers in Human Behavior for their continuous support and collaboration while working on this special issue. The guest editors would also like to thank the authors and reviewers without whom this special issue would not be possible.

This special issue was in part supported by funding from Australian Research Council (DP210100060, DP220101209), Economic and Social Research Council of the United Kingdom (ES/S015701/1), and Jacobs Foundation (CELLA 2 CERES) awarded to the first author. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Australian Research Council, Economic and Social Research Council of the United Kingdom, or Jacobs Foundation.

This special issue was funded in part by the National Science Foundation (DRL-1661036, DRL-1713110, DRL-2100320), the Wisconsin Alumni Research Foundation, and the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison. The opinions, findings, and conclusions do not reflect the views of the funding agencies, cooperating institutions, or other individuals.

6 References

Aguilar, S. J., Karabenick, S. A., Teasley, S. D., & Baek, C. (2021). Associations between learning analytics dashboard exposure and motivation and self-regulated learning. *Computers & Education*, *162*, 104085. https://doi.org/10.1016/j.compedu.2020.104085

Alexandron, G., Ruipérez-Valiente, J. A., Chen, Z., Muñoz-Merino, P. J., & Pritchard, D. E. (2017). Copying@Scale: Using Harvesting Accounts for Collecting Correct Answers in a MOOC. *Computers & Education*, *108*, 96–114. https://doi.org/10.1016/j.compedu.2017.01.015

Arnold, K. E., & Pistilli, M. D. (2012). Course signals at Purdue: Using learning analytics to increase student success. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 267–270. https://doi.org/10.1145/2330601.2330666

Azevedo, R., & Gašević, D. (2019). Analyzing Multimodal Multichannel Data about Self-Regulated Learning with Advanced Learning Technologies: Issues and Challenges. *Computers in Human Behavior*, *96*, 207–210. https://doi.org/10.1016/j.chb.2019.03.025

Baker, R. S., Gašević, D., & Karumbaiah, S. (2021). Four paradigms in learning analytics: Why paradigm convergence matters. *Computers and Education: Artificial Intelligence*, *2*, 100021. https://doi.org/10.1016/j.caeai.2021.100021

Barthakur, A., Kovanovic, V., Joksimovic, S., Siemens, G., Richey, M., & Dawson, S. (2021). Assessing program-level learning strategies in MOOCs. *Computers in Human Behavior*, *117*, 106674. https://doi.org/10.1016/j.chb.2020.106674

Beck, R. J., Skinner, W. F., & Schwabrow, L. A. (2013). A study of sustainable assessment theory in higher education tutorials. *Assessment & Evaluation in Higher Education*, *38*(3), 326–348. https://doi.org/10.1080/02602938.2011.630978

Bodily, R., & Verbert, K. (2017). Review of Research on Student-Facing Learning Analytics Dashboards and Educational Recommender Systems. *IEEE Transactions on Learning Technologies*, *10*(4), 405–418. https://doi.org/10.1109/TLT.2017.2740172

Carter, M., & Egliston, B. (2021). What are the risks of virtual reality data? Learning analytics, algorithmic bias and a fantasy of perfect data. *New Media & Society*, In press. https://doi.org/10.1177/14614448211012794

Chaturapruek, S., Dee, T. S., Johari, R., Kizilcec, R. F., & Stevens, M. L. (2018). How a Data-driven Course Planning Tool Affects College Students' GPA: Evidence from Two Field Experiments. *Proceedings of the Fifth Annual ACM Conference on Learning at Scale*, 63:1-63:10. https://doi.org/10.1145/3231644.3231668

Chen, B., Knight, S., & Wise, A. F. (2018). Critical Issues in Designing and Implementing Temporal Analytics. *Journal of Learning Analytics*, *5*(1), 1–9. https://doi.org/10.18608/jla.2018.53.1

Chouldechova, A., & Roth, A. (2018). The Frontiers of Fairness in Machine Learning. *ArXiv:1810.08810*.

Clarebout, G., Elen, J., Collazo, N. A. J., Lust, G., & Jiang, L. (2013). Metacognition and the Use of Tools. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 187–195). Springer New York. http://link.springer.com/chapter/10.1007/978-1-4419-5546-3_13

Coghlan, S., Miller, T., & Paterson, J. (2021). Good proctor or "big brother"? Ethics of online exam supervision technologies. *Philosophy & Technology*, *34*(4), 1581–1606.

Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction*, *4*(4), 253–278. https://doi.org/10.1007/BF01099821

Crocker, L., & Algina, J. (2009). *Introduction to classical and modern test theory*. Cengage Learning.

Desmarais, M. C., & Baker, R. S. (2012). A Review of Recent Advances in Learner and Skill Modeling in Intelligent Learning Environments. *User Modeling and User-Adapted Interaction*, *22*(1–2), 9–38. https://doi.org/10.1007/s11257-011-9106-8

D'Mello, S. K. (2017). Emotional Learning Analytics. In C. Lang, G. Siemens, W. Alyssa, & D. Gašević (Eds.), *Handbook of Learning Analytics & Educational Data Mining* (First edition). Societz for Learning Analytics Research.

Dowell, N. M. M., & Poquet, O. (2021). SCIP: Combining group communication and interpersonal positioning to identify emergent roles in scaled digital environments. *Computers in Human Behavior*, *119*, 106709. https://doi.org/10.1016/j.chb.2021.106709

Fan, Y., van der Graaf, J., Lim, L., Kilgour, J., Raković, M., Singh, S., Moore, J., Molenaar, I., Bannert, M., & Gašević, D. (2022). Towards improving the validity of measurement of self-regulated learning based on trace data. *Metacognition and Learning*, in press.

Ferguson, R., Hoel, T., Scheffel, M., & Drachsler, H. (2016). Guest Editorial: Ethics and Privacy in Learning Analytics. *Journal of Learning Analytics*, *3*(1), 5–15. https://doi.org/10.18608/jla.2016.31.2

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231–235). Lawrence Erlbaum Associate.

Gardner, J., Brooks, C., & Baker, R. S. J. d. (2019). Evaluating the Fairness of Predictive Student Models Through Slicing Analysis. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 225–234.

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*(1), 64–71. https://doi.org/10.1007/s11528-014-0822-x

Gašević, D., Mirriahi, N., Dawson, S., & Joksimović, S. (2017). Effects of instructional conditions and experience on the adoption of a learning tool. *Computers in Human Behavior*, *67*, 207–220. https://doi.org/10.1016/j.chb.2016.10.026

Herder, T., Swiecki, Z., Fougt, S. S., Tamborg, A. L., Allsopp, B. B., Shaffer, D. W., & Misfeldt, M. (2018). Supporting teachers' intervention in students' virtual collaboration using a network based model. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, 21–25.

Ifenthaler, D., & Greiff, S. (2021). Leveraging Learning Analytics for Assessment and Feedback. In J. Liebowitz (Ed.), *Online Learning Analytics* (pp. 1–18). CRC Press.

Joksimović, S., Manataki, A., Gašević, D., Dawson, S., Kovanović, V., & de Kereki, I. F. (2016). Translating Network Position into Performance: Importance of Centrality in Different Network Configurations. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, 314–323. https://doi.org/10.1145/2883851.2883928

Jones, L. V., & Appelbaum, M. I. (1989). Psychometric methods. *Annual Review of Psychology*, *40*(1), 23–43.

Jovanović, J., Saqr, M., Joksimović, S., & Gašević, D. (2021). Students matter the most in learning analytics: The effects of internal and instructional conditions in predicting academic success. *Computers & Education*, *172*, 104251. https://doi.org/10.1016/j.compedu.2021.104251

Kane, M. T. (2013). Validating the Interpretations and Uses of Test Scores. *Journal of Educational Measurement*, *50*(1), 1–73. https://doi.org/10.1111/jedm.12000

Kharbat, F. F., & Abu Daabes, A. S. (2021). E-proctored exams during the COVID-19 pandemic: A close understanding. *Education and Information Technologies*, *26*(6), 6589–6605. https://doi.org/10.1007/s10639-021-10458-7

Kitto, K., & Knight, S. (2019). Practical ethics for building learning analytics. *British Journal of Educational Technology*, *50*(6), 2855–2870. https://doi.org/10.1111/bjet.12868

Knight, S., Buckingham Shum, S., & Littleton, K. (2013). Epistemology, Pedagogy, Assessment and Learning Analytics. *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, 75–84. https://doi.org/10.1145/2460296.2460312

Kollom, K., Tammets, K., Scheffel, M., Tsai, Y.-S., Jivet, I., Muñoz-Merino, P. J., Moreno-Marcos, P. M., Whitelock-Wainwright, A., Calleja, A. R., Gasevic, D., Kloos, C. D., Drachsler, H., & Ley, T. (2021). A four-country cross-case analysis of academic staff expectations about learning analytics in higher education. *The Internet and Higher Education*, *49*, 100788. https://doi.org/10.1016/j.iheduc.2020.100788

Lang, C., Siemens, G., Wise, A., Gašević, D., & Merceron, A. (Eds.). (2022). *Handbook of learning analytics* (Second). Society for Learning Analytics and Research.

Lim, L.-A., Gentili, S., Pardo, A., Kovanović, V., Whitelock-Wainwright, A., Gašević, D., & Dawson, S. (2021). What changes, and for whom? A study of the impact of learning analytics-based process feedback in a large course. *Learning and Instruction*, *72*, 101202. https://doi.org/10.1016/j.learninstruc.2019.04.003

Liu, M., Kitto, K., & Buckingham Shum, S. (2021). Combining factor analysis with writing analytics for the formative assessment of written reflection. *Computers in Human Behavior*, *120*, 106733. https://doi.org/10.1016/j.chb.2021.106733

Lodge, J., & Lewis, M. (2012). Pigeon pecks and mouse clicks: Putting the learning back into learning analytics. *In Proceedings of The29th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education (Ascilite 2012)*, 560–564.

Lonn, S., Aguilar, S. J., & Teasley, S. D. (2015). Investigating student motivation in the context of a learning analytics intervention during a summer bridge program. *Computers in Human Behavior*, *47*, 90–97. https://doi.org/10.1016/j.chb.2014.07.013

Macfadyen, L. P., Dawson, S., Pardo, A., & Gasevic, D. (2014). Embracing big data in complex educational systems: The learning analytics imperative and the policy challenge. *Research & Practice in Assessment*, *9*(2), 17–28.

Martínez-Maldonado, R., Yan, L., Deppeler, J., Phillips, M., & Gašević, D. (2022). Classroom Analytics: Telling Stories About Learning Spaces Using Sensor Data. In E. Gil, Y. Mor, Y. Dimitriadis, & C. Köppe (Eds.), *Hybrid Learning Spaces* (pp. 185–203). Springer International Publishing. https://doi.org/10.1007/978-3-030-88520-5_11

Matcha, W., Gašević, D., Uzir, N. A., Jovanović, J., Pardo, A., Lim, L., Maldonado-Mahauad, J., Gentili, S., Pérez-Sanagustín, M., & Tsai, Y.-S. (2020). Analytics of Learning Strategies: Role of Course Design and Delivery Modality. *Journal of Learning Analytics*, *7*(2), 45–71. https://doi.org/10.18608/jla.2020.72.3

Mayfield, E., Madaio, M., Prabhumoye, D., Gerritsen, B. M., Dixon-Román, E., & Black, A. W. (2019). Equity beyond bias in language technologies for education. In *Proceedings of the 14th Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 444–460).

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2019). A Survey on Bias and Fairness in Machine Learning. *ArXiv:1908.09635*.

Messick, S. (1994). Validity of Psychological Assessment: Validation of Inferences from Persons' Responses and Performances as Scientific Inquiry into Score Meaning. *ETS Research Report Series*, *1994*(2), i–28. https://doi.org/10.1002/j.2333-8504.1994.tb01618.x

Messick, S. (1995). Standards of Validity and the Validity of Standards in Performance Asessment. *Educational Measurement: Issues and Practice*, *14*(4), 5–8. https://doi.org/10.1111/j.1745-3992.1995.tb00881.x

Milligan, S. (2018). Methodological foundations for the measurement of learning in learning analytics. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, 466–470.

Milligan, S. (2020). Standards for Developing Assessments of Learning Using Process Data. In M. Bearman, P. Dawson, R. Ajjawi, J. Tai, & D. Boud (Eds.), *Re-imagining University Assessment in a Digital World* (pp. 179–192). Springer International Publishing. https://doi.org/10.1007/978-3-030-41956-1_13

Milligan, S., & Griffin, P. (2016). Understanding Learning and Learning Design in MOOCs: A Measurement-Based Interpretation. *Journal of Learning Analytics*, *3*(2), 88–115. https://doi.org/10.18608/jla.2016.32.5

Misiejuk, K., Wasson, B., & Egelandsdal, K. (2021). Using learning analytics to understand student perceptions of peer feedback. *Computers in Human Behavior*, *117*, 106658. https://doi.org/10.1016/j.chb.2020.106658

Mislevy, R. J., Haertel, G., Riconscente, M., Rutstein, D. W., & Ziker, C. (2017). Evidence-Centered Assessment Design. In *Assessing Model-Based Reasoning using Evidence-Centered Design* (pp. 19–24). Springer, Cham. https://doi.org/10.1007/978-3-319-52246-3_3

Nicolay, B., Krieger, F., Stadler, M., Gobert, J., & Greiff, S. (2021). Lost in transition – Learning analytics on the transfer from knowledge acquisition to knowledge application in complex problem solving. *Computers in Human Behavior*, *115*, 106594. https://doi.org/10.1016/j.chb.2020.106594

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.

Pardo, A. (2018). A feedback model for data-rich learning experiences. *Assessment & Evaluation in Higher Education*, *43*(3), 428–438. https://doi.org/10.1080/02602938.2017.1356905

Peters, H., Kyngdon, A., & Stillwell, D. (2021). Construction and validation of a game-based intelligence assessment in minecraft. *Computers in Human Behavior*, *119*, 106701. https://doi.org/10.1016/j.chb.2021.106701

Poquet, O., & Jovanovic, J. (2020). Intergroup and interpersonal forum positioning in shared-thread and post-reply networks. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 187–196). Association for Computing Machinery. https://doi.org/10.1145/3375462.3375533

Prinsloo, P., & Slade, S. (2018). Mapping responsible learning analytics: A critical proposal. In B. H. Khan, J. R. Corbeil, & M. E. Corbeil (Eds.), *Responsible Analytics and Data Mining in Education: Global Perspectives on Quality, Support, and Decision Making* (First, pp. 63–79). Routledge.

Rogers, T., Gašević, D., & Dawson, S. (2016). Learning analytics and the imperative for theory driven research. In C. Haythornthwaite, R. Andrews, J. Fransma, & E. Meyers (Eds.), *The SAGE Handbook of E-Learning Research, 2nd edition* (pp. 232–250). SAGE Publications Ltd.

Rolim, V., Ferreira, R., Lins, R. D., & Gašević, D. (2019). A network-based analytic approach to uncovering the relationship between social and cognitive presences in communities of inquiry. *The Internet and Higher Education*, *42*, 53–65. https://doi.org/10.1016/j.iheduc.2019.05.001

Rowe, E., Almeda, M. V., Asbell-Clarke, J., Scruggs, R., Baker, R., Bardar, E., & Gasca, S. (2021). Assessing implicit computational thinking in Zoombinis puzzle gameplay. *Computers in Human Behavior*, *120*, 106707. https://doi.org/10.1016/j.chb.2021.106707

Ruipérez-Valiente, J. A., Jaramillo-Morillo, D., Joksimović, S., Kovanović, V., Muñoz-Merino, P. J., & Gašević, D. (2021). Data-driven detection and characterization of communities of accounts collaborating in MOOCs. *Future Generation Computer Systems*, *125*, 590–603.

Saint, J., Fan, Y., Pardo, A., & Gaševic, D. (2022). Temporally-focused analytics of self-regulated learning: A systematic review of literature. *Computers & Education: Artificial Intelligence*, in press.

Selwyn, N. (2020). Re-imagining 'Learning Analytics'… a case for starting again? *The Internet and Higher Education*, *46*, 100745.

Selwyn, N., O'Neill, C., Smith, G., Andrejevic, M., & Gu, X. (2021). A necessary evil? The rise of online exam proctoring in Australian universities. *Media International Australia*. https://doi.org/10.1177/1329878X211005862

Sha, L., Rakovic, M., Whitelock-Wainwright, A., Carroll, D., Yew, V. M., Gasevic, D., & Chen, G. (2021). Assessing Algorithmic Fairness in Automatic Classifiers of Educational Forum Posts. *Proceedings of the 22nd International Conference on Artificial Intelligence in Education*, 381–394. https://doi.org/10.1007/978-3-030-78292-4_31

Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A Tutorial on Epistemic Network Analysis: Analyzing the Structure of Connections in Cognitive, Social, and Interaction Data. *Journal of Learning Analytics*, *3*(3), 9–45. https://doi.org/10.18608/jla.2016.33.3

Sharma, K., & Giannakos, M. (2020). Multimodal data capabilities for learning: What can multimodal data tell us about learning? *British Journal of Educational Technology*, *51*(5), 1450–1484. https://doi.org/10.1111/bjet.12993

Shute, V. J., & Rahimi, S. (2021). Stealth assessment of creativity in a physics video game. *Computers in Human Behavior*, *116*, 106647. https://doi.org/10.1016/j.chb.2020.106647

Siemens, G. (2013). Learning Analytics The Emergence of a Discipline. *American Behavioral Scientist*, *57*(10), 1380–1400. https://doi.org/10.1177/0002764213498851

Stadler, M., Hofer, S., & Greiff, S. (2020). First among equals: Log data indicates ability differences despite equal scores. *Computers in Human Behavior*, *111*, 106442. https://doi.org/10.1016/j.chb.2020.106442

Taras, M. (2008). Assessment for learning: Sectarian divisions of terminology and concepts. *Journal of Further and Higher Education*, *32*(4), 389–397.

Tsai, Y.-S., Moreno-Marcos, P. M., Jivet, I., Scheffel, M., Tammets, K., Kollom, K., & Gašević, D. (2018). The SHEILA Framework: Informing Institutional Strategies and Policy Processes of Learning Analytics. *Journal of Learning Analytics*, *5*(3), 5–20. https://doi.org/10.18608/jla.2018.53.2

Tsai, Y.-S., Whitelock-Wainwright, A., & Gašević, D. (2021). More than figures on your laptop:(Dis) trustful implementation of learning analytics. *Journal of Learning Analytics*, *8*(3), 81–100.

Van Der Graaf, J., Lim, L., Fan, Y., Kilgour, J., Moore, J., Bannert, M., Gasevic, D., & Molenaar, I. (2021). Do Instrumentation Tools Capture Self-Regulated Learning? *Proceedings of The11th International Learning Analytics and Knowledge Conference*, 438–448. https://doi.org/10.1145/3448139.3448181

VanLehn, K. (2008). Intelligent tutoring systems for continuous, embedded assessment. *The Future of Assessment: Shaping Teaching and Learning*, 113–138.

Winne, P. H. (2006). How Software Technologies Can Improve Research on Learning and Bolster School Reform. *Educational Psychologist*, *41*(1), 5–17. https://doi.org/10.1207/s15326985ep4101_3

Winne, P. H. (2020). Construct and consequential validity for learning analytics based on trace data. *Computers in Human Behavior*, *112*, 106457. https://doi.org/10.1016/j.chb.2020.106457

Wise, A., Knight, S., & Shum, S. B. (2021). Collaborative Learning Analytics. In U. Cress, C. Rosé, A. F. Wise, & J. Oshima (Eds.), *International Handbook of Computer-Supported Collaborative Learning* (pp. 425–443). Springer International Publishing. https://doi.org/10.1007/978-3-030-65291-3_23

Wise, A., & Shaffer, D. W. (2015). Why Theory Matters More than Ever in the Age of Big Data. *Journal of Learning Analytics*, *2*(2), 5–13. https://doi.org/10.18608/jla.2015.22.2

Worsley, M., Martinez-Maldonado, R., & D'Angelo, C. (2021). A New Era in Multimodal Learning Analytics: Twelve Core Commitments to Ground and Grow MMLA. *Journal of Learning Analytics*, *8*(3), 10–27. https://doi.org/10.18608/jla.2021.7361

Zhang, S., Bergner, Y., DiTrapani, J., & Jeon, M. (2021). Modeling the interaction between resilience and ability in assessments with allowances for multiple attempts. *Computers in Human Behavior*, *122*, 106847. https://doi.org/10.1016/j.chb.2021.106847

Table 1. Themes, authors, and brief description of the contributions included in the special issue

Theme	Authors	Description
Analytics for assessment	Abhinava Barthakur, Vitomir Kovanovic, Srecko Joksimovic, George Siemens, Michael Richey, Shane Dawson	The study proposes a learning analytic approach for longitudinal assessment of learning strategies based on latent class analysis of online trace data collected within several MOOCs that are part of a professional development program. The study identified three program-level strategies that were significantly associated with outcomes. The study also found a significant effect of MOOC design on the level of student engagement.
	Heinrich Peters, Andrew Kyngdon, David Stillwell	The study proposes the use of video game Minecraft™ for assessment of intelligence. Intelligence was measured for pattern completion, mental rotation, and spatial construction. The results showed moderate reliability with Rasch models; factorial validity with separate factors for pattern completion and spatial construction tasks, but not for mental rotation. Trace data were very predictive of performance in the Minecraft™ tests; trace data were also predictive of performance on conventional tests.
	Elizabeth Rowe, Ma Victoria Almeda, Jodi Asbell-Clarke, Richard Scruggs, Ryan Baker, ErinBardar, Santiago Gasca	The study examines the use of machine learning classifiers for automatic detection of implicit practices in computational think. The study analyzed trace data about behavior while playing the game called Zoombinis™. The study showed a good reliability of automatic detectors in comparison to that by expert coders. The external validity of the automatically detected practices was confirmed through strong correlations with post-assessment scores.
	Nia M.M. Dowell, Oleksandra Poquet	The study proposed an analytic approach for assessment of emergent socio-cognitive roles that learners adopt in online social interactions. The approach is based on a combination of group communication analysis and social network analysis. The approach was applied to a dataset collected in a MOOC and found five emergent socio-cognitive roles that learners took while interacting with their peers.
Analytics of assessment	Matthias Stadler, Sarah Hofer, Samuel Greiff	The paper makes use of trace data about student testing behavior to check whether test-taking behavior is a good indicator of tested ability. The study used structural equation modeling (SEM) in the context of complex problem solving assessment to show that both time-on-task and the count of interactions were significant predictors of students' GPA. However, when intelligence was added to SEM, time-on-task and count of interaction become almost negligible predictors.
	Bjorn Nicolay, Florian Krieger, Matthias Stadler, Janice Gobert, Samuel Greif	The paper made use of trace data to check whether learners transfer knowledge acquired to knowledge application during a complex problem solving task. The study showed that many learners were not able to transfer their knowledge from acquisition to application. The number of learners who were unable to make this translation was associated with the complexity of assessment items.
	Kamila Misiejuk, Barbara Wasson, Kjetil Egelandsdal	The study investigated students' reactions to peer assessment using epistemic network analysis. The results unveiled that students value specificity, justification, and constructiveness while kindness was not as much appreciated in peer assessment. The study also revealed

		differences between students who did and did not found peer assessment useful.
	Susu Zhang, Yoav Bergner, Jack DiTrapani, Minjeong Jeon	The study proposed a novel approach to modeling the interaction between resilience (i.e., persistence in multiple attempts in computer-based assessments) and ability in assessments. The results showed resilience to affect performance scores, and such can be a threat to the validity of summative assessments.
Validity	Philip H. Winne	The conceptual paper analyzes the notion of validity and reliability in learning analytics based on trace data. The paper emphasizes the critical role of theory in assuring validity in learning analytics and the consideration of dynamic nature of events about learning in contrast to the static nature of conventional measures.
	Valerie J. Shute and Seyedahmad Rahimi	The study reports on stealth assessment of creativity in a physics video game. The study showed an external validity of the stealth assessments through the significant correlations with the external measures of creativity, in-game performance, game enjoyment, and learning of physics content.
	Ming Liu, Kirsty Kitto, and Simon Buckingham Shum	The study reports on the findings of a model for automated formative assessment of written reflection. The study validated the model by using confirmatory factor analysis of textual features of written reflections from two different datasets. The writing context was found to have a significant impact on the validity of the proposed model.