Embracing trustworthiness and authenticity in the validation of learning analytics systems

Haastrecht, M.A.N. van; Brinkhuis, M.J.S.; Peichl, J.; Remmele, B.; Spruit, M.

Embracing Trustworthiness and Authenticity in the Validation of Learning Analytics Systems

Max van Haastrecht*
Leiden University
Leiden, The Netherlands
m.a.n.van.haastrecht@liacs.leidenuniv.nl

Matthieu Brinkhuis
Utrecht University
Utrecht, The Netherlands
m.j.s.brinkhuis@uu.nl

Jessica Peichl
University of Education Freiburg
Freiburg, Germany
jessica.peichl@ph-freiburg.de

Bernd Remmele
University of Education Freiburg
Freiburg, Germany
bernd.remmele@ph-freiburg.de

Marco Spruit
Leiden University Medical Center
Leiden, The Netherlands
Leiden University
Leiden, The Netherlands
m.r.spruit@lumc.nl

ABSTRACT

Learning analytics sits in the middle space between learning theory and data analytics. The inherent diversity of learning analytics manifests itself in an epistemology that strikes a balance between positivism and interpretivism, and knowledge that is sourced from theory and practice. In this paper, we argue that validation approaches for learning analytics systems should be cognisant of these diverse foundations. Through a systematic review of learning analytics validation research, we find that there is currently an over-reliance on positivistic validity criteria. Researchers tend to ignore interpretivistic criteria such as trustworthiness and authenticity. In the 38 papers we analysed, researchers covered positivistic validity criteria 221 times, whereas interpretivistic criteria were mentioned 37 times. We motivate that learning analytics can only move forward with holistic validation strategies that incorporate "thick descriptions" of educational experiences. We conclude by outlining a planned validation study using argument-based validation, which we believe will yield meaningful insights by considering a diverse spectrum of validity criteria.

CCS CONCEPTS

• **Applied computing** → **Learning management systems**; • **Human-centered computing** → *Collaborative and social computing design and evaluation methods.*

KEYWORDS

learning analytics, validation, trustworthiness, authenticity, interpretivism

*Corresponding author.

LAK 2023, March 13–17, 2023, Arlington, TX, USA

ACM ISBN 978-1-4503-9865-7/23/03.
https://doi.org/10.1145/3576050.3576060

ACM Reference Format:
Max van Haastrecht, Matthieu Brinkhuis, Jessica Peichl, Bernd Remmele, and Marco Spruit. 2023. Embracing Trustworthiness and Authenticity in the Validation of Learning Analytics Systems. In *LAK23: 13th International Learning Analytics and Knowledge Conference (LAK 2023), March 13–17, 2023, Arlington, TX, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3576050.3576060

1 INTRODUCTION

In a recent survey among learning analytics experts [16], validity was ranked as the third-most important theme relating to the future of learning analytics, behind power (i.e., control over data) and pedagogy. Ferguson et al. [16] state that validation approaches should always take "context into account when reporting results". Recognising that each instructional context is different is seen by Gašević et al. [20] as a prerequisite for an acceptable validation strategy. Kitto et al. [34] agree, arguing that validation must address both positivistic (e.g., performance metrics) and interpretivistic (e.g., student experience) elements. They conclude that "work on developing new validation criteria that emphasise learning outcomes" is vital. This conclusion is in agreement with the experts in Ferguson et al. [16], who state that "research in this space should be tied to pedagogical outcomes."

Thus, validation is a critical topic for learning analytics research. There is agreement that validation should go beyond performance metrics and that an additional emphasis on learning outcomes would help to yield a contextualised approach. Yet, there is little consensus on which validity criteria are essential in learning analytics research. In a recent special issue on the potential links between learning analytics and educational assessment, Gašević et al. [21] raised the concern that "existing learning analytic methods do not meet all of the criteria" for validation we encounter in educational assessment. However, Gašević et al. [21] do not discuss to which criteria they are referring. The learning analytics literature lacks an in-depth analysis of the validity criteria that are currently in use and the criteria that deserve emphasis. We will address this gap in this paper.

With the previous paragraphs in mind, we formulate the following main research question and sub-questions:

- **RQ**: Which validity criteria should be considered in a contextualised validation strategy for learning analytics systems?
  - **RQa**: Which validity criteria have emerged in the learning analytics domain that emphasise learning outcomes?
  - **RQb**: How has learning analytics validation research incorporated interpretivistic perspectives that recognise contextual differences?

Through an analysis of the epistemological foundations of learning analytics (Section 2) and a systematic review of the learning analytics validation literature (Section 3), we will construct an overview of emerging validity criteria to answer **RQa**. An in-depth analysis of our systematic review results (Section 4) will help us in answering **RQb**. We discuss the implications for our main research question in Section 5 and conclude in Section 6.

2 BACKGROUND: EPISTEMOLOGY AS A FOUNDATION FOR VALIDATION

How we approach validation depends on our underlying epistemology, specifically relating to our view on the concept of truth. A purely interpretivist researcher will attach little value to performance metrics when validating since they reject the concept of objective truth in social contexts. Similarly, positivist researchers are unlikely to engage in what Geertz [22] termed "thick description" of social contexts, as they believe in the generalisability of more efficiently obtainable quantitative evidence. We posit that learning analytics epistemology is positioned in the middle space between interpretivism and positivism. In this section, we will provide further intuition for this observation and motivate that the axis of truth is not the only epistemological axis relevant to building a solid foundation for validation.

Pragmatism is one of the cornerstones of today's learning analytics literature. As envisioned by Dewey [9], pragmatism takes a moderate position in the interpretivism versus positivism debate. Kuhn [37] describes the scientific process as "a process whose successive stages are characterised by an increasingly detailed and refined understanding of nature." A process of moving "from primitive beginnings," yet not "towards anything." This contradicts the positivist view that the scientific method enables us to consistently hone in on truths and thereby expand our knowledge. Dewey [10] avoids the term knowledge altogether, preferring "warranted assertability." This phrase connects the past (warranted) and the future (assertability). Dewey's pragmatism, therefore, blends views that aim to build from a common past (interpretivism) with those that aim to move towards a common future (positivism).

However, the axis of truth is not the only relevant epistemological axis when laying the foundations for validation. Pragmatists claim that "our conception of some given thing is bound up in our understanding of its practical application" [36]. Not only a definition of what constitutes knowledge is crucial, but also a consideration of possible sources of knowledge. Pragmatism posits that practical use should be the primary source of knowledge, which juxtaposes it with rationalism which states that theoretical reasoning is the summum bonum when it comes to knowledge gathering. Wise et al. [59] propose a similar classification regarding learning analytics design knowledge. They state that design knowledge can originate from the design process, which is guided by theory, and from the implementation process, which is coupled with the introduction of learning analytics in the learning environment.

Dewey helped develop a version of pragmatism, known as transactionalism, that emphasises contextual interactions as a vital source of knowledge [11]. Transactionalism merges ideas from pragmatism and constructivism, with Dewey's version of pragmatism being considered "as the most important precursor for social constructivism" [47]. Social constructivism is the variant of constructivism most often encountered in learning analytics research today. Social constructivists argue "that learners arrive at what they know mainly through participating in the social practices of a learning environment" [60]. Social constructivism focuses on meaningful interactions in authentic contexts. However, in today's world, many educational interactions involve technological assistance. Although there is a role for social constructivism in technology-enhanced learning [60], its focus on social interactions as the primary source of knowledge makes it ill-suited to assess the consequences of today's socio-technical systems. Siemens [51] aimed to solve this issue with connectivism.

Connectivism is perhaps the philosophical stance most closely associated with learning analytics. Connectivism is similar to social constructivism, but it reserves an explicit place for "learning that occurs outside of people (i.e. learning that is stored and manipulated by technology)" [51]. Connectivism, like learning analytics itself, states that theory is a valid source of knowledge. This brings us, finally, to the place that learning analytics epistemology occupies within the epistemological plane of Figure 1. We propose that learning analytics epistemology is positioned in the middle space between positivism and interpretivism on the axis of truth, but also in the middle space between theory and practice on the knowledge source axis.

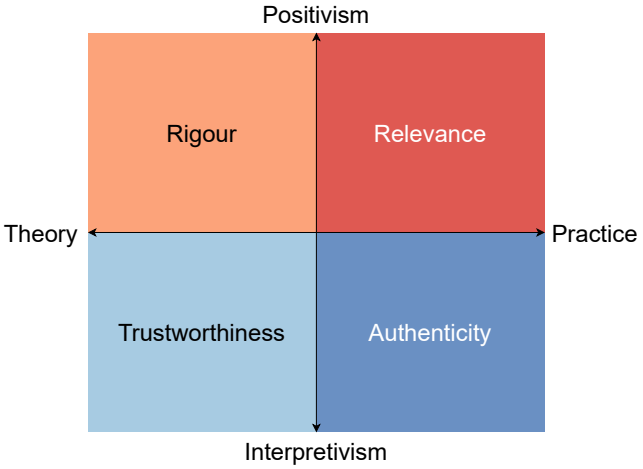

Figure 1: Our epistemological plane of validity, divided into four quadrants. Learning analytics occupies the middle space between positivism and interpretivism (the truth axis), and the middle space between theory and practice (the knowledge source axis).

Figure 1 introduces the overarching terms we use within this paper to refer to the four quadrants created by the axes of our

epistemological plane. In the positivistic tradition, it is common to distinguish between rigour and relevance in research [25]. Rigour is connected to theory as a source of knowledge and can be "achieved by appropriately applying existing foundations and methodologies" [25]. Research is relevant when it addresses "the problems faced and the opportunities afforded by the interaction of people, organisations, and information technology" [25]. On the side of interpretivism, Guba [24] proposed the concept of trustworthiness as a parallel to rigour. Lincoln and Guba [38] later introduced authenticity as a more practice-oriented validity conceptualisation, noting that "conventional criteria refer only to methodology and ignore the influence of context."

Figure 1 only presents four overarching validity quadrants, providing an incomplete answer to **RQa** on emerging validity criteria. Many more criteria are considered in the learning analytics literature, each with its own place within the epistemological plane. To investigate which criteria are considered and whether specific areas of the epistemological plane are underrepresented, we conducted a systematic review of the learning analytics validation literature.

3 METHODOLOGY

For our systematic review, we queried three databases: ACM Digital Library, Web of Science, and PubMed. We searched for all papers with abstracts containing the phrase 'learning analytics' and either 'validation' or 'validity'. After deduplicating the query results, 83 papers remained. Of these papers, 21 formed the initial set of inclusions after excluding work that did not discuss validation or was unrelated to the field of learning analytics (as defined by SoLAR [53]). For each of these 21 included papers, we scanned all the references and citations to find potential new inclusions. This process is known as 'snowballing' and is a recommended step in systematic review methodologies [56]. The snowballing phase resulted in a further 17 inclusions, meaning our final set comprised 38 papers.

Before proceeding to analyse our inclusions, we identified four papers which would allow us to construct a holistic set of potential validity criteria. We first looked towards educational measurement (sometimes referred to as educational assessment). Educational measurement is a field where validity considerations naturally take centre stage, and several learning analytics researchers have argued that we should strengthen the bond with this field [21]. The argument-based validation approach of Kane [31] has been influential in the educational measurement and learning analytics fields in recent years [12, 41]. Kane [31] stresses the importance of addressing traditional validity criteria such as rigour, construct validity, content validity, and criterion validity. However, Kane's framework also recognises that theoretical considerations alone are insufficient, and that validation must investigate how results are used in practice. Kane captures this idea in the concept of consequential validity.

The fields of design science and information systems offer a second source of inspiration in the validity considerations made by learning analytics researchers. Mingers and Standing [42] provide an extensive overview of the validation literature in these fields, while highlighting the importance of the interpretivistic perspective. The criteria external validity (sometimes termed generalisability), internal validity, reliability, replicability, and statistical validity occupy the rigour quadrant. Mingers and Standing [42] additionally propose consistency (relevance quadrant) and elegance (authenticity quadrant) criteria.

Our third external source of validity terminology is the seminal interpretivistic work of Lincoln and Guba [38]. Their paper introduced the concept of authenticity as a counterbalance to trustworthiness. Lincoln and Guba [38] discuss various dimensions of trustworthiness that parallel positivistic criteria: confirmability (related to replicability and content validity), credibility (internal validity), dependability (reliability), and transferability (external validity). They additionally discuss several dimensions of authenticity, but we select to include authenticity as a single criterion in this paper as this is generally how the construct is viewed in learning analytics research. Lastly, Lincoln and Guba [38] introduce fairness as a vital consideration during validation.

Finally, certain validity considerations are quite unique to the learning analytics field. To provide sufficient coverage of these validity criteria, we looked towards the work of Ali et al. [2]. They propose a diverse selection of validity criteria covering the relevance quadrant (relevance, actionability, understandability, usability, and usefulness) and the authenticity quadrant (meaningfulness and parsimony/simplicity).

4 RESULTS

Figure 2 depicts the assembled validity criteria within their respective quadrants. Criteria are positioned according to how they are defined and treated in the literature, thereby acting as a Learning Analytics Validation Assistant (LAVA). Researchers can use LAVA to determine whether the validity criteria they are considering are sufficient and appropriate for their epistemological stance. A criterion's quadrant is determined by how it is defined in one of the four core papers mentioned in the previous section. The exact placement of a criterion within a quadrant should not be interpreted as an indisputable truth. Rather, we positioned criteria relative to each other based on how they were treated and measured in the learning analytics literature.

Figure 2 additionally visualises the prevalence of the validity criteria in our included papers. In 38 inclusions, a total of 258 validity criteria were discussed. Criteria in the rigour quadrant were mentioned 146 times (56.6%), in the relevance quadrant 75 times (29.1%), in the trustworthiness quadrant 11 times (4.3%), and in the authenticity quadrant 26 times (10.1%). Hence, researchers covered positivistic criteria 221 times, whereas interpretivistic criteria were mentioned only 37 times.

Statistical validity and external validity are the criteria mentioned most often within our inclusions. For statistical validity, we noticed that most papers focus on statistical significance, whereas Saqr and López-Pernas [49] point out that researchers should additionally consider effect size. External validity is another problematic criterion within learning analytics research. Of the 27 times external validity was mentioned in one of our inclusions, 24 times the authors concluded that the external validity of their study was lacking. We observed a similar pattern with the interpretivistic counterpart to external validity: transferability. Of the three times transferability was considered, the authors stated on two occasions

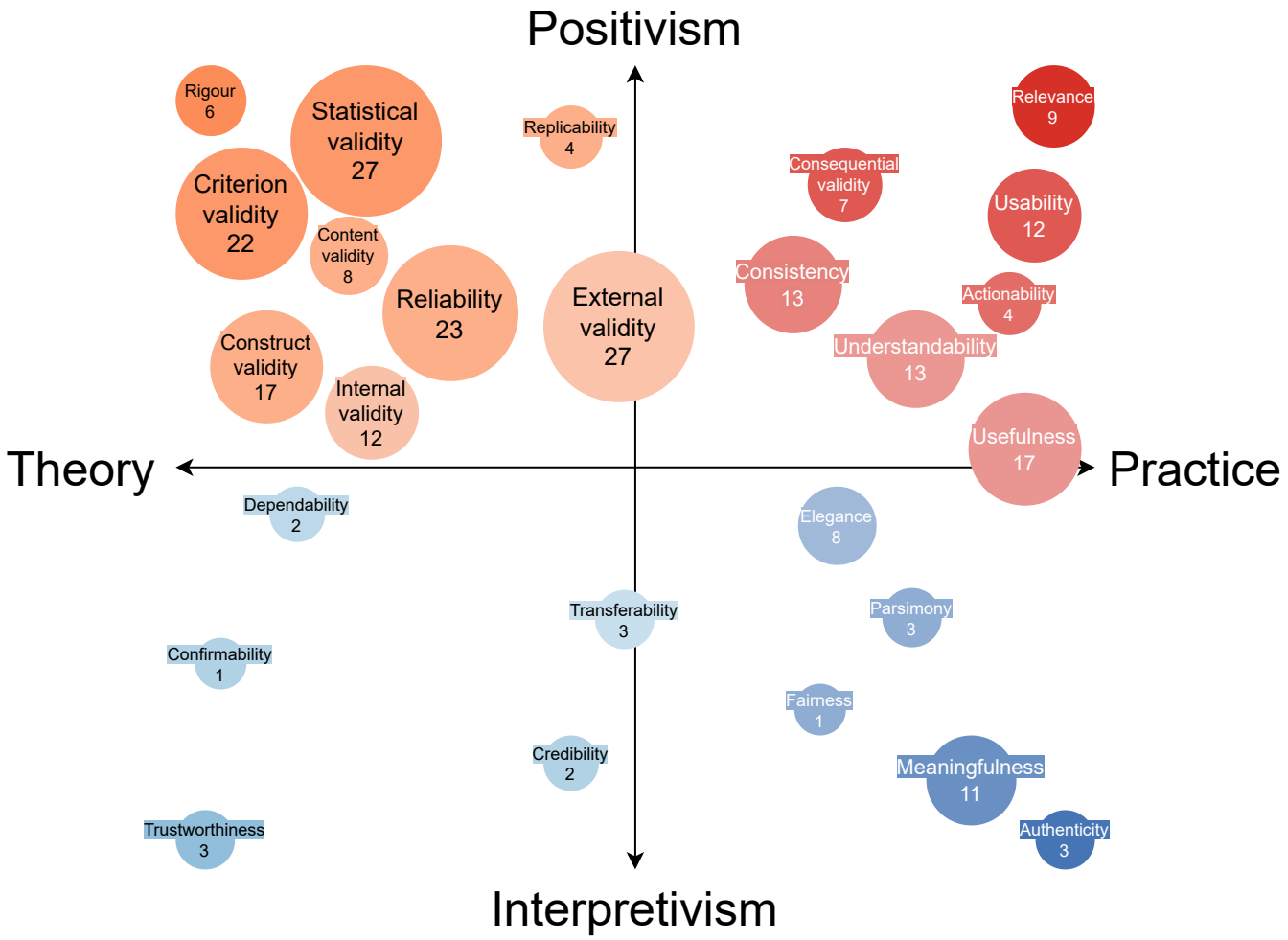

Figure 2: The Learning Analytics Validation Assistant (LAVA), depicting the prevalence of validity criteria observed in the learning analytics literature. Terms are positioned along the two axes (truth: positivism versus interpretivism; knowledge source: theory versus practice) of our epistemological plane.

that more work was necessary to assess the transferability of their results.

Figure 2 provides an answer to **RQa**: Which validity criteria have emerged in the learning analytics domain that emphasise learning outcomes? Criteria on the 'practice' half of the diagram relate to outcomes of the learning process. Criteria positioned on the extreme right of the theory-practice axis correspond to an advanced internalisation of learning analytics outcomes. Learning analytics researchers evidently attach importance to relevant, usable, actionable, and useful solutions. Additionally, several papers recognised that authentic, meaningful learning experiences are not simply a luxury, but a goal to strive for.

Table 1 lists the combinations of validity criteria quadrants observed in our inclusions. Four out of 38 papers covered criteria from all four quadrants. Papers tended to consider criteria from at least two quadrants, with only one inclusion not covering a criterion from the rigour quadrant. Conversely, the criteria in the trustworthiness quadrant, along with parsimony, authenticity, and fairness, are mentioned least often. Many of these criteria can only be assessed through "thick descriptions" of social contexts [22], possibly pointing to barriers to engaging in such activities within learning analytics research. Moreover, although meaningfulness was discussed in 11 papers, only one of these papers conducted qualitative interviews during validation. All other papers used either quantitative data analysis or structured questionnaires in their evaluation. Concerning **RQb**, we can conclude that although interpretivistic validity criteria are considered in learning analytics research, their treatment is often too superficial to provide in-depth insight into contextual learning experiences.

Table 1: Combinations of the four validity criteria quadrants (rigour, relevance, trustworthiness, and authenticity) observed in the 38 inclusions of our systematic review, sorted by number of related inclusions. Only observed combinations are listed.

Quadrant combination	Related inclusions
Rigour, relevance	[4, 7, 13, 14, 17, 18, 23, 27, 50, 54]
Rigour, relevance, authenticity	[3, 29, 32, 43–45, 49, 57, 61, 63]
Rigour	[5, 30, 39, 40, 46]
Rigour, authenticity	[8, 15, 33, 35, 52]
Rigour, relevance, trustworthiness, authenticity	[2, 55, 59, 62]
Rigour, relevance, trustworthiness	[6, 48, 58]
Relevance	[19]

5 DISCUSSION

Our results lead to three main findings related to the learning analytics validation literature, which we will cover in this section.

5.1 Troubling External Validity

Learning analytics researchers seem to have a troubling relationship with external validity. Together with statistical validity, external validity was the criterion mentioned most often in our inclusions. Yet, 24 out of the 27 papers that mention external validity conclude that there are limitations to the generalisability of their results. At times, the limited scale of studies is listed as the cause for generalisability concerns (e.g., [7, 54, 59, 62]). Elsewhere, researchers provide a general warning that more research is necessary should one want to generalise the results (e.g., [14, 33, 48]). Transferability, the interpretivistic parallel of external validity, suffers from the same issue. Researchers state that results could be transferred to other contexts, but that more research is required to confirm this claim [2, 6].

The reader should not interpret the previous paragraph as a critique of the cited research. If there are limitations to the generalisability of findings, these should be mentioned. However, we should avoid a situation in the learning analytics field where generalisability becomes an afterthought that can always be left for future work. External validity and transferability are valued validity criteria that should guide learning analytics research a priori, not a posteriori. Replication studies that aim to understand the validity of learning analytics solutions in new contexts should receive more attention.

5.2 A Need for Thick Descriptions

We noted in Section 4 that even papers that recognise interpretive validity criteria (e.g., meaningfulness) often resort to quantitative methods during validation. Geertz [22] believes that the analysis of social culture and context requires qualitative methods "in search of meaning" rather than quantitative methods "in search of law." In other words, we require "thick descriptions" of the educational contexts being considered in learning analytics research. Thick descriptions that cannot be obtained through data analysis or questionnaires, but that require qualitative methods.

The advantages of using qualitative methods go beyond a deeper understanding of the educational context. As Guba [24] recognises, "to determine the extent to which transferability is probable, one needs to know a great deal about both the transferring and receiving contexts." Guba [24] states that thick descriptions are essential if we wish to achieve transferable results. Thus, thick descriptions provide deeper insight into interpretivistic validity criteria and concurrently act as a catalyst in facilitating generalisable learning analytics research. Researchers looking to produce more generalisable results will benefit from employing qualitative research methods such as qualitative interviews and action research.

5.3 The Potential of Argument-Based Validation

To conclude this section, we will discuss a validation approach uniquely suited to facilitate the diverse validity criteria and research methods covered in this paper: argument-based validation. Kane [31] originally introduced this approach in the educational measurement field. Gašević et al. [21] argue that learning analytics research can profit from the vast validity experience within educational measurement and psychological assessment, and argument-based validation has started to see use within the learning analytics domain [12, 41].

In general, research uses inferences to make warranted claims based on data. Argument-based validation proceeds by constructing arguments to provide evidence for the assertability of these claims. Once evidence has been assembled in structured arguments, we assess the validity of the overall inference chain. The benefit of this approach is that it gives a balance of flexibility and structure, allowing researchers to recognise "legitimately diverse arguments" [1] while avoiding the open-ended nature of validation. The original framework of Kane [31] has been extended to allow for an increased focus on practical consequences [26] and to explicitly address fairness in artificial intelligence (AI) enhanced assessments [28]. Argument-based validation is a promising avenue for learning analytics researchers looking to address diverse validity criteria and produce rigorous, relevant, trustworthy, and authentic results.

6 CONCLUSION AND FUTURE WORK

Within this paper, we have investigated which validity criteria should be considered in a contextualised validation strategy for learning analytics systems. We proceeded by first analysing the epistemological foundations of learning analytics research, concluding that learning analytics epistemology is positioned in the middle space between positivism and interpretivism and between theory and practice. We then conducted a systematic review to uncover which types of validity criteria are employed by learning

analytics researchers. We visualised the results to create a Learning Analytics Validation Assistant (LAVA).

We uncovered an over-reliance on positivistic criteria. Interpretivistic criteria that were covered (e.g., meaningfulness), were often investigated using quantitative rather than qualitative methods. In Section 5, we analysed the LAVA results and delineated a need for more focus on "thick descriptions" of educational experiences. Such thick descriptions help to foster a deeper understanding of the context being studied and can act as a catalyst in facilitating generalisable research.

In future work, we will apply our LAVA insights within an educational research project. As suggested in Section 5.1, we intend to employ an argument-based validation approach incorporating diverse arguments and validity criteria. We recognise that we are bound to encounter limitations in our future work and want to stress that no single approach can function as a validation panacea. Nevertheless, we believe that LAVA can stimulate researchers to evaluate whether their validity criteria are sufficient and appropriate for their epistemological stance.

ACKNOWLEDGMENTS

This work was made possible with funding from the European Union's Horizon 2020 research and innovation programme, under grant agreement No. 883588 (GEIGER), and funding from a European Commission Erasmus+ project, under project code 2022-1-DE02-KA220-VET-000087221 (MECyS). The opinions expressed and arguments employed herein do not necessarily reflect the official views of the funding bodies.

REFERENCES

[1] Camilla Addey, Bryan Maddox, and Bruno D. Zumbo. 2020. Assembled Validity: Rethinking Kane's Argument-Based Approach in the Context of International Large-Scale Assessments (ILSAs). *Assessment in Education: Principles, Policy & Practice* 27, 6 (2020), 588–606.

[2] Liaqat Ali, Marek Hatala, Dragan Gašević, and Jelena Jovanović. 2012. A Qualitative Evaluation of Evolution of a Learning Analytics Tool. *Computers & Education* 58, 1 (2012), 470–489. https://doi.org/10.1016/j.compedu.2011.08.030

[3] Cristina Alonso-Fernández, Ana R. Cano, Antonio Calvo-Morata, Manuel Freire, Iván Martínez-Ortiz, and Baltasar Fernández-Manjón. 2019. Lessons Learned Applying Learning Analytics to Assess Serious Games. *Computers in Human Behavior* 99 (2019), 301–309.

[4] Norman B. Berman and Anthony R. Artino. 2018. Development and Initial Validation of an Online Engagement Metric Using Virtual Patients. *BMC Medical Education* 18, 1 (2018), 213. https://doi.org/10.1186/s12909-018-1322-z

[5] Robyn Bitner, Nguyen-Thinh Le, and Niels Pinkwart. 2020. A Concurrent Validity Approach for EEG-Based Feature Classification Algorithms in Learning Analytics. In *Proceedings of the 12th International Conference on Computational Collective Intelligence (ICCCI'20)*. Springer-Verlag, Da Nang, Vietnam, 568–580. https://doi.org/10.1007/978-3-030-63007-2_44

[6] Juan Pedro Cerro Martínez, Montse Guitert Catasús, and Teresa Romeu Fontanillas. 2020. Impact of Using Learning Analytics in Asynchronous Online Discussions in Higher Education. *International Journal of Educational Technology in Higher Education* 17, 1 (2020), 39.

[7] Julián Chaparro-Peláez, Santiago Iglesias-Pradas, Francisco J. Rodríguez-Sedano, and Emiliano Acquila-Natale. 2020. Extraction, Processing and Visualization of Peer Assessment Data in Moodle. *Applied Sciences* 10, 1 (2020), 163. https://doi.org/10.3390/app10010163

[8] Pankaj Chejara, Luis P. Prieto, Adolfo Ruiz-Calleja, María Jesús Rodríguez-Triana, Shashi Kant Shankar, and Reet Kasepalu. 2021. EFAR-MMLA: An Evaluation Framework to Assess and Report Generalizability of Machine Learning Models in MMLA. *Sensors* 21, 8 (2021), 2863.

[9] John Dewey. 1931. *Philosophy and Civilization.* Minton, Balch & Company, New York, NY, USA.

[10] John Dewey. 1938. *Logic: The Theory of Inquiry.* Henry Holt & Company, New York, NY, USA.

[11] John Dewey and Arthur Fisher Bentley. 1949. *Knowing and the Known.* Beacon Press, Boston, MA, USA.

[12] Kerrie A. Douglas, Hillary E. Merzdorf, Nathan M. Hicks, Muhammad Ihsanulhaq Sarfraz, and Peter Bermel. 2020. Challenges to Assessing Motivation in MOOC Learners: An Application of an Argument-Based Approach. *Computers & Education* 150 (2020), 103829.

[13] Raphael A. Dourado, Rodrigo Lins Rodrigues, Nivan Ferreira, Rafael Ferreira Mello, Alex Sandro Gomes, and Katrien Verbert. 2021. A Teacher-facing Learning Analytics Dashboard for Process-oriented Feedback in Online Learning. In *Proceedings of the 11th International Learning Analytics and Knowledge Conference (LAK'21)*. ACM, Irvine, CA, USA, 482–489. https://doi.org/10.1145/3448139.3448187

[14] Tomas Effenberger and Radek Pelánek. 2021. Validity and Reliability of Student Models for Problem-Solving Activities. In *Proceedings of the 11th International Learning Analytics and Knowledge Conference (LAK'21)*. ACM, Irvine, CA, USA, 1–11. https://doi.org/10.1145/3448139.3448140

[15] Yizhou Fan, Lyn Lim, Joep van der Graaf, Jonathan Kilgour, Mladen Raković, Johanna Moore, Inge Molenaar, Maria Bannert, and Dragan Gašević. 2022. Improving the Measurement of Self-Regulated Learning Using Multi-Channel Data. *Metacognition and Learning* 17, 3 (2022), 1025–1055.

[16] Rebecca Ferguson, Doug Clow, Dai Griffiths, and Andrew Brasher. 2019. Moving Forward with Learning Analytics: Expert Views. *Journal of Learning Analytics* 6, 3 (2019), 43–59. https://doi.org/10.18608/jla.2019.63.8

[17] Ed Fincham, Alexander Whitelock-Wainwright, Vitomir Kovanović, Srećko Joksimović, Jan-Paul van Staalduinen, and Dragan Gašević. 2019. Counting Clicks Is Not Enough: Validating a Theorized Model of Engagement in Learning Analytics. In *Proceedings of the 9th International Learning Analytics and Knowledge Conference (LAK'19)*. ACM, Tempe, AZ, USA, 501–510. https://doi.org/10.1145/3303772.3303775

[18] Joy Galaige, Geraldine Torrisi-Steele, Sebastian Binnewies, and Kewen Wang. 2018. What Is Important in Student-Facing Learning Analytics? A User-Centered Design Approach. In *Proceedings of the 22nd Pacific Asia Conference on Information Systems (PACIS'18)*. AIS, Yokohoma, Japan, 1248–1261.

[19] David Gañán, Santi Caballé, Robert Clarisó, Jordi Conesa, and David Bañeres. 2017. ICT-FLAG: A Web-Based e-Assessment Platform Featuring Learning Analytics and Gamification. *International Journal of Web Information Systems* 13, 1 (2017), 25–54.

[20] Dragan Gašević, Shane Dawson, Tim Rogers, and Danijela Gasevic. 2016. Learning Analytics Should Not Promote One Size Fits All: The Effects of Instructional Conditions in Predicting Academic Success. *The Internet and Higher Education* 28 (2016), 68–84.

[21] Dragan Gašević, Samuel Greiff, and David Williamson Shaffer. 2022. Towards Strengthening Links between Learning Analytics and Assessment: Challenges and Potentials of a Promising New Bond. *Computers in Human Behavior* 134 (2022), 107304. https://doi.org/10.1016/j.chb.2022.107304

[22] Clifford Geertz. 1973. Thick Description: Toward an Interpretive Theory of Culture. In *The Interpretation Of Cultures*. Basic Books, New York, NY, USA, 3–30.

[23] Michail N. Giannakos, Konstantinos Chorianopoulos, and Nikos Chrisochoides. 2015. Making Sense of Video Analytics: Lessons Learned from Clickstream Interactions, Attitudes, and Learning Outcome in a Video-Assisted Course. *International Review of Research in Open and Distributed Learning* 16, 1 (2015), 260–283. https://doi.org/10.19173/irrodl.v16i1.1976

[24] Egon G. Guba. 1981. Criteria for Assessing the Trustworthiness of Naturalistic Inquiries. *ECTJ* 29, 2 (1981), 75–91. https://doi.org/10.1007/BF02766777

[25] Alan R. Hevner, Salvatore T. March, Jinsoo Park, and Sudha Ram. 2004. Design Science in Information Systems Research. *MIS Quarterly* 28, 1 (2004), 75–105. https://doi.org/10.2307/25148625

[26] Dorien Hopster-den Otter, Saskia Wools, Theo J. H. M. Eggen, and Bernard P. Veldkamp. 2019. A General Framework for the Validation of Embedded Formative Assessment. *Journal of Educational Measurement* 56, 4 (2019), 715–732. https://doi.org/10.1111/jedm.12234

[27] Joel A. Howell, Lynne D. Roberts, and Vincent O. Mancini. 2018. Learning Analytics Messages: Impact of Grade, Sender, Comparative Information and Message Style on Student Affect and Academic Resilience. *Computers in Human Behavior* 89 (2018), 8–15. https://doi.org/10.1016/j.chb.2018.07.021

[28] A. Corinne Huggins-Manley, Brandon M. Booth, and Sidney K. D'Mello. 2022. Toward Argument-Based Fairness with an Application to AI-Enhanced Educational Assessments. *Journal of Educational Measurement* 59, 3 (2022), 362–388. https://doi.org/10.1111/jedm.12334

[29] Dirk Ifenthaler and Chathuranga Widanapathirana. 2014. Development and Validation of a Learning Analytics Framework: Two Case Studies Using Support Vector Machines. *Technology, Knowledge and Learning* 19, 1 (2014), 221–240. https://doi.org/10.1007/s10758-014-9226-4

[30] Il-Hyun Jo, Dongho Kim, and Meehyun Yoon. 2014. Analyzing the Log Patterns of Adult Learners in LMS Using Learning Analytics. In *Proceedings of the 4th International Learning Analytics and Knowledge Conference (LAK '14)*. ACM, Indianapolis, IN, USA, 183–187.

[31] Michael T. Kane. 2013. Validating the Interpretations and Uses of Test Scores. *Journal of Educational Measurement* 50, 1 (2013), 1–73.

[32] Tobias Kärner, Julia Warwas, and Stephan Schumann. 2021. A Learning Analytics Approach to Address Heterogeneity in the Classroom: The Teachers' Diagnostic Support System. *Technology, Knowledge and Learning* 26, 1 (2021), 31–52. https://doi.org/10.1007/s10758-020-09448-4

[33] Dongho Kim, Yeonjeong Park, Meehyun Yoon, and Il-Hyun Jo. 2016. Toward Evidence-Based Learning Analytics: Using Proxy Variables to Improve Asynchronous Online Discussion Environments. *The Internet and Higher Education* 30 (2016), 30–43. https://doi.org/10.1016/j.iheduc.2016.03.002

[34] Kirsty Kitto, Simon Buckingham Shum, and Andrew Gibson. 2018. Embracing Imperfection in Learning Analytics. In *Proceedings of the 8th International Learning Analytics and Knowledge Conference (LAK '18)*. ACM, Sydney, Australia, 451–460. https://doi.org/10.1145/3170358.3170413

[35] René F. Kizilcec, Mar Pérez-Sanagustín, and Jorge J. Maldonado. 2017. Self-Regulated Learning Strategies Predict Learner Behavior and Goal Attainment in Massive Open Online Courses. *Computers & Education* 104 (2017), 18–33. https://doi.org/10.1016/j.compedu.2016.10.001

[36] Simon Knight, Simon Buckingham Shum, and Karen Littleton. 2014. Epistemology, Assessment, Pedagogy: Where Learning Meets Analytics in the Middle Space. *Journal of Learning Analytics* 1, 2 (2014), 23–47–23–47. https://doi.org/10.18608/jla.2014.12.3

[37] Thomas S. Kuhn. 1962. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago, IL, USA.

[38] Yvonna S. Lincoln and Egon G. Guba. 1986. But Is It Rigorous? Trustworthiness and Authenticity in Naturalistic Evaluation. *New Directions for Program Evaluation* 1986, 30 (1986), 73–84. https://doi.org/10.1002/ev.1427

[39] Jorge Maldonado-Mahauad, Mar Pérez-Sanagustín, René F. Kizilcec, Nicolás Morales, and Jorge Munoz-Gama. 2018. Mining Theory-Based Patterns from Big Data: Identifying Self-Regulated Learning Strategies in Massive Open Online Courses. *Computers in Human Behavior* 80 (2018), 179–196.

[40] Wannisa Matcha, Dragan Gašević, Jelena Jovanović, Nora'ayu Ahmad Uzir, Chris W Oliver, Andrew Murray, and Danijela Gasevic. 2020. Analytics of Learning Strategies: The Association with the Personality Traits. In *Proceedings of the 10th International Learning Analytics and Knowledge Conference (LAK '20)*. ACM, Frankfurt, Germany, 151–160. https://doi.org/10.1145/3375462.3375534

[41] Sandra K. Milligan. 2018. Methodological Foundations for the Measurement of Learning in Learning Analytics. In *Proceedings of the 8th International Learning Analytics and Knowledge Conference (LAK '18)*. ACM, Sydney, Australia, 466–470. https://doi.org/10.1145/3170358.3170391

[42] John Mingers and Craig Standing. 2020. A Framework for Validating Information Systems Research Based on a Pluralist Account of Truth and Correctness. *Journal of the Association for Information Systems* 21, 1 (2020), 6. https://doi.org/10.17705/1jais.00594

[43] Sergio Muñoz, Enrique Sánchez, and Carlos A. Iglesias. 2020. An Emotion-Aware Learning Analytics System Based on Semantic Task Automation. *Electronics* 9, 8 (2020), 1194. https://doi.org/10.3390/electronics9081194

[44] Abelardo Pardo, Robert A. Ellis, and Rafael A. Calvo. 2015. Combining Observational and Experiential Data to Inform the Redesign of Learning Activities. In *Proceedings of the 5th International Learning Analytics and Knowledge Conference (LAK'15)*. ACM, Poughkeepsie, NY, USA, 305–309.

[45] Yeonjeong Park and Il-Hyun Jo. 2019. Factors That Affect the Success of Learning Analytics Dashboards. *Educational Technology Research and Development* 67, 6 (2019), 1547–1571. https://doi.org/10.1007/s11423-019-09693-0

[46] Alain Prat and Warren J. Code. 2021. WeBWorK Log Files as a Rich Source of Data on Student Homework Behaviours. *International Journal of Mathematical Education in Science and Technology* 52, 10 (2021), 1540–1556. https://doi.org/10.1080/0020739X.2020.1782492

[47] Kersten Reich. 2007. Interactive Constructivism in Education. *Education and Culture* 23, 1 (2007), 7–26.

[48] Mohammed Saqr, Uno Fors, and Jalal Nouri. 2018. Using Social Network Analysis to Understand Online Problem-Based Learning and Predict Performance. *PLOS ONE* 13, 9 (2018), e0203590. https://doi.org/10.1371/journal.pone.0203590

[49] Mohammed Saqr and Sonsoles López-Pernas. 2021. Modelling Diffusion in Computer-Supported Collaborative Learning: A Large Scale Learning Analytics Study. *International Journal of Computer-Supported Collaborative Learning* 16, 4 (2021), 441–483. https://doi.org/10.1007/s11412-021-09356-4

[50] Mohammed Saqr, Olga Viberg, and Henriikka Vartiainen. 2020. Capturing the Participation and Social Dimensions of Computer-Supported Collaborative Learning through Social Network Analysis: Which Method and Measures Matter? *International Journal of Computer-Supported Collaborative Learning* 15, 2 (2020), 227–248. https://doi.org/10.1007/s11412-020-09322-6

[51] George Siemens. 2004. Connectivism: A Learning Theory for the Digital Age. *International Journal of Instructional Technology and Distance Learning* 2, 1 (2004), 1–8.

[52] Tanmay Sinha, Patrick Jermann, Nan Li, and Pierre Dillenbourg. 2014. Your Click Decides Your Fate: Inferring Information Processing and Attrition Behavior from MOOC Video Clickstream Interactions. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*. ACL, Doha, Qatar, 3–14. https://doi.org/10.3115/v1/W14-4102

[53] Society for Learning Analytics Research (SoLAR). 2022. What Is Learning Analytics? https://www.solaresearch.org/about/what-is-learning-analytics/.

[54] Bernardo Tabuenca, Marco Kalz, Hendrik Drachsler, and Marcus Specht. 2015. Time Will Tell: The Role of Mobile Learning Analytics in Self-Regulated Learning. *Computers & Education* 89 (2015), 53–74. https://doi.org/10.1016/j.compedu.2015.08.004

[55] Natercia Valle, Pavlo Antonenko, Denis Valle, Max Sommer, Anne Corinne Huggins-Manley, Kara Dawson, Dongho Kim, and Benjamin Baiser. 2021. Predict or Describe? How Learning Analytics Dashboard Design Influences Motivation and Statistics Anxiety in an Online Statistics Course. *Educational Technology Research and Development* 69, 3 (2021), 1405–1431. https://doi.org/10.1007/s11423-021-09998-z

[56] Max van Haastrecht, Injy Sarhan, Bilge Yigit Ozkan, Matthieu Brinkhuis, and Marco Spruit. 2021. SYMBALS: A Systematic Review Methodology Blending Active Learning and Snowballing. *Frontiers in Research Metrics and Analytics* 6 (2021), 685591. https://doi.org/10.3389/frma.2021.685591

[57] Alexander Whitelock-Wainwright, Dragan Gašević, Yi-Shan Tsai, Hendrik Drachsler, Maren Scheffel, Pedro J. Muñoz-Merino, Kairit Tammets, and Carlos Delgado Kloos. 2020. Assessing the Validity of a Learning Analytics Expectation Instrument: A Multinational Study. *Journal of Computer Assisted Learning* 36, 2 (2020), 209–240. https://doi.org/10.1111/jcal.12401

[58] Philip H. Winne. 2020. Construct and Consequential Validity for Learning Analytics Based on Trace Data. *Computers in Human Behavior* 112 (2020), 106457. https://doi.org/10.1016/j.chb.2020.106457

[59] Alyssa Friend Wise, Jovita Maria Vytasek, Simone Hausknecht, and Yuting Zhao. 2016. Developing Learning Analytics Design Knowledge in the "Middle Space": The Student Tuning Model and Align Design Framework for Learning Analytics Use. *Online Learning* 20, 2 (2016), 155–182.

[60] Younghee Woo and Thomas C. Reeves. 2007. Meaningful Interaction in Web-Based Learning: A Social Constructivist Interpretation. *The Internet and Higher Education* 10, 1 (2007), 15–25. https://doi.org/10.1016/j.iheduc.2006.10.005

[61] Dan Ye and Svoboda Pennisi. 2022. Using Trace Data to Enhance Students' Self-Regulation: A Learning Analytics Perspective. *The Internet and Higher Education* 54 (2022), 100855. https://doi.org/10.1016/j.iheduc.2022.100855

[62] Mina Yoo and Sung-Hee Jin. 2020. Development and Evaluation of Learning Analytics Dashboards to Support Online Discussion Activities. *Educational Technology & Society* 23, 2 (2020), 1–18.

[63] Lanqin Zheng, Jiayu Niu, and Lu Zhong. 2022. Effects of a Learning Analytics-Based Real-Time Feedback Approach on Knowledge Elaboration, Knowledge Convergence, Interactive Relationships and Group Performance in CSCL. *British Journal of Educational Technology* 53, 1 (2022), 130–149.