



## OPEN ACCESS

EDITED AND REVIEWED BY  
Gavin T. L. Brown,  
The University of Auckland, New Zealand

\*CORRESPONDENCE  
Kelly Anne Young  
✉ youngka@unisa.ac.za

RECEIVED 15 April 2025  
ACCEPTED 16 April 2025  
PUBLISHED 29 April 2025

CITATION  
Archer E, Young KA, Grover R and Khalil M  
(2025) Editorial: Educational evaluation in the  
age of artificial intelligence: challenges and  
innovations. *Front. Educ.* 10:1612274.  
doi: 10.3389/feduc.2025.1612274

COPYRIGHT  
© 2025 Archer, Young, Grover and Khalil. This  
is an open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with these  
terms.

# Editorial: Educational evaluation in the age of artificial intelligence: challenges and innovations

Elizabeth Archer<sup>1</sup>, Kelly Anne Young<sup>2\*</sup>, Raman Grover<sup>3</sup> and  
Mohammad Khalil<sup>4</sup>

<sup>1</sup>University of the Western Cape, Bellville, South Africa, <sup>2</sup>University of South Africa, Pretoria, South Africa, <sup>3</sup>University of British Columbia, Vancouver, BC, Canada, <sup>4</sup>University of Bergen, Bergen, Norway

## KEYWORDS

educational evaluation, artificial intelligence, challenges, innovations, assessment

## Editorial on the Research Topic

### Educational evaluation in the age of artificial intelligence: challenges and innovations

The rapid evolution of artificial intelligence in education is reshaping how learning is measured and redefining the theoretical and methodological foundations of assessment. This Research Topic creates a dedicated forum for rigorous academic debate on the transformative potential of AI in educational evaluation. By examining innovations in automated scoring, dynamic assessment creation, data-driven analytics, and early detection, the contributions in this Research Topic illuminate both the promise and the challenges inherent in this digital transformation.

Wang et al. set the stage with their study on improving automated scoring of prosody in oral reading fluency using a deep learning algorithm. Employing a carefully engineered selection of prosodic features and a sophisticated neural network, the authors demonstrate enhanced cross-domain accuracy in fluency assessment. Their work reinforces the importance of feature selection and model interpretability and offers a robust framework that could recalibrate traditional reading assessments by integrating acoustic and linguistic cues.

Building on the theme of automation, Twabu and Nakene-Mginqi present an innovative design thinking approach to developing an AI-driven auto-marking and grading system tailored for higher education in South Africa. Their iterative, user-centered methodology integrates principles from design thinking with advanced machine-learning techniques to streamline the grading process. Their system exemplifies how digital transformation can enhance operational efficiency while maintaining academic rigor by reducing lecturer workload and ensuring consistent, bias-mitigated feedback.

Hamady et al. contribute a dynamic perspective by integrating GeoGebra and Moodle to create multi-version interactive assessments. Their approach leverages the interactive capabilities of GeoGebra to generate animated, unique test items while using Moodle's robust learning management framework to administer assessments securely. This method not only combats issues of academic dishonesty through randomized test forms but also enriches the student experience by offering varied, context-sensitive problem sets. The study highlights the potential of integrating interactive tools to elevate assessments' engagement and validity.

Mehrabi et al. further advance our understanding of AI-enhanced evaluation by refining performance factor analysis. Their research introduces an attention mechanism to integrate detailed skill profiles and item similarity, providing a nuanced perspective on student learning trajectories. The authors' methodology underscores the importance of capturing the interdependencies between emergent and internalized skills, offering an analytical model that could improve the predictive accuracy of educational outcomes and inform personalized learning strategies.

Complementing these approaches, Borna et al. explore using clickstream data to predict student performance. By applying advanced machine learning models to digital interaction data, they demonstrate how behavioral analytics can serve as an early-warning system. Their work provides valuable insights into the correlations between online engagement and academic success, suggesting that data-driven interventions could significantly enhance the responsiveness of educational environments.

Foldnes et al. address the crucial issue of early detection with their study on identifying struggling readers using gameplay data and machine learning. The authors use a rich dataset derived from a pedagogical literacy game to develop predictive models that pinpoint students at risk of reading difficulties at school entry. Their research highlights the potential for timely, non-intrusive diagnostics that can trigger early interventions, thereby mitigating long-term academic challenges and supporting equitable learning outcomes.

Rounding out the Research Topic, Ateeq et al. critically examine the broader implications of AI in education by focusing on academic integrity and the shift toward holistic assessment. Their work interrogates the ethical and theoretical underpinnings of AI-driven evaluation, advocating for assessment practices that balance automated efficiency with human oversight. By considering issues such as algorithmic bias, data privacy, and the erosion of traditional evaluative measures, their study calls for robust policy frameworks and comprehensive reforms that ensure reliability and fairness in the digital age.

In conclusion, the studies presented in this Research Topic collectively chart a transformative course for educational evaluation. They illustrate how AI can automate and enrich assessment practices through innovative methodologies and data-driven insights while raising critical questions about validity, reliability, and ethical implementation. Future research must delve deeper into these challenges—developing more nuanced theoretical frameworks, ensuring equitable access, and continuously evaluating the long-term impacts of digital transformation on education. The field can harness AI's full

potential to create robust, adaptive, and just educational environments for the digital age by addressing these specific challenges and exploring detailed future directions.

Together, these contributions provide a comprehensive vision of how AI is poised to reshape educational evaluation, urging researchers and practitioners alike to engage with both its transformative benefits and the imperative for thoughtful, critical oversight.

## Author contributions

EA: Writing – original draft. KY: Writing – review & editing. RG: Writing – review & editing. MK: Writing – review & editing.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

## Generative AI statement

The author(s) declare that Gen AI was used in the creation of this manuscript. This editorial includes sections where initial drafts were generated using ChatGPT-4.5 (March 2025). These sections were subsequently reviewed, critically edited, and validated by the human authors to ensure academic rigor and adherence to ethical standards. The author(s) also assessed and addressed potential biases inherent in the AI-generated content. The final content, conclusions, and assertions in this paper are the sole responsibility of the human authors.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.