InsPIRE Insight

Mapping the landscape: Generative AI in higher education assessment (2020-2024) - a scoping review

By

- S. H. S. Ng (InsPIRE; CRADLE)
- H. Y. Chan (InsPIRE)
- J. H. K. Wong (Office of Information, Knowledge and Library Services)
- L. Sam (Office of Information, Knowledge and Library Services)
- A. J. Privitera (Science of Learning in Education Centre, NIE)

Nanyang Technological University
Singapore

KEY IMPLICATIONS

This scoping review to map pioneering efforts studying Generative AI in higher education assessment reveals:

- Text-based GAI tools, particularly ChatGPT, dominate assessment research, indicating a critical need for educators to explore multimodal tools that can support diverse assessment formats.
- While GAI offers significant potential to enhance efficiency in assessment design and feedback, a "human-in-the-loop" approach is essential to mitigate risks associated with factual inaccuracies and "hallucinations".
- There is a notable absence of research addressing equity and access in GAI-enhanced assessment, requiring urgent attention to ensure that integration does not exacerbate existing educational inequalities.

BACKGROUND

The emergence of Generative Artificial Intelligence (GAI) tools marks a paradigm shift in higher education, presenting disruptive capabilities that challenge traditional assessment practices. While there have been a number of reviews synthesising research on GAI in education, many were limited by excluding grey literature (such as conference papers and preprints). This omission is significant as grey literature often captures the most pioneering work in this rapidly evolving field. To address this gap, a comprehensive mapping of the landscape was necessary to guide informed decision-making and address the fragmentation of knowledge.

FOCUS OF INITIATIVE

This scoping review aimed to provide a comprehensive global overview of how GAI has been

used in higher education assessment between June 2020 and February 2024. The study investigated which tools were utilised, the academic disciplines involved, and the specific areas of assessment targeted (e.g., design, feedback, academic integrity). It sought to identify current practices and critical research gaps to support practitioners and researchers in navigating the integration of GAI into assessment strategies.

KEY OUTCOMES

- **Dominance of ChatGPT:** OpenAI's ChatGPT or other GPT models were the most frequently investigated tools, appearing in approximately 88% of the included documents, while non-text-based tools remain underutilised.
- **Disciplinary Divide:** STEM disciplines accounted for 52% of the studies, with Computer Science leading the exploration; conversely, non-STEM disciplines represented only 32%, suggesting a potential lag in these fields.
- **Assessment Areas:** The majority of studies focused on perceptions of GAI use (26 studies), followed by assessment design (18 studies) and academic integrity (16 studies).
- **Efficiency vs. Reliability:** In assessment design, GAI was found to efficiently create high-quality questions. However, for grading and feedback, results were mixed; while often comprehensive and motivating, GAI output sometimes contained inaccuracies, necessitating educator oversight.

SIGNIFICANCE OF OUTCOMES

- **Implications for practice:** Educators should view GAI as a tool to augment and scale feedback and grading processes rather than fully automate them. A "human-in-the-loop" approach is strictly necessary, where educators retain final oversight to ensure quality and pedagogic value.
- **Implications for policy:** Institutions should move away from an "arms race" focused solely on prohibition and detection, especially as classifiers fail to reliably detect AI-generated text. Instead, policies should foster a culture of ethical engagement and AI literacy, establishing clear guidelines to support responsible implementation.
- **Proposed follow-up activities:**
    - Conduct longitudinal studies to understand the long-term impact of GAI on student learning, critical thinking, and skill development.
    - Urgently investigate equity and access considerations to ensure that GAI integration does not disadvantage specific student populations.
    - Explore effective strategies for integrating human judgement with GAI capabilities to maximise the benefits of both.

PARTICIPANTS/SCOPE

The review included 68 documents (36 articles, 1 review, 8 preprints, and 23 conference papers) published between 1 June 2020 and 29 February 2024. The studies represented higher education institutions globally, with high concentrations in the United States, United

Kingdom, and India.

METHODOLOGY/APPROACH

The study employed a scoping review methodology guided by Arksey and O'Malley's (2005) five-stage framework. Comprehensive searches were conducted across ERIC, ProQuest, Scopus, and Web of Science databases to include both peer-reviewed and grey literature. Screening and data extraction were performed by three independent reviewers using the Covidence platform to ensure rigour.

References

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology, 8*(1), 19-32.

Link to full article: https://doi.org/10.1080/10494820.2026.2614079

Contact S. H. S. Ng at sally.ngsh@ntu.edu.sg for more information about the initiative.