A Tale of Two Assessments: Navigating Generative AI Integration in Authentic Assessment Design

Ramalakshmi Vaidhiyanathan, Cardiff University, United Kingdom

The European Conference on Education 2026
Official Conference Proceedings

Abstract

The study presents a comparative analysis of two approaches to authentic assessment design, highlighting how scaffolding and generative AI integration influence student learning in computing education. In the first approach, students engaged in structured coursework supported by predefined requirements and starter code. A custom AI chatbot was provided to enhance comprehension and guide processes while deliberately restricting code generation to uphold academic integrity. This structured design fostered strong student engagement, reduced instructor intervention, and supported consistent understanding of expectations. Although some evidence suggested limited AI-assisted code generation, the controlled scope mitigated the risk of overreliance on automated solutions. In contrast, the second approach adopted an open-ended use of generative AI, permitting students to employ AI tools as virtual clients to create requirement specifications and to use AI-generated code with declaration. This high-autonomy model unintentionally elevated student anxiety and increased requests for instructor support. Students faced difficulties determining appropriate requirements to demonstrate learning outcomes and struggled to connect AI outputs to underlying design principles. Instead of empowering learners, unrestricted AI use resulted in decision paralysis and conceptual confusion. Together, these findings underscore that autonomy without adequate scaffolding can overwhelm learners engaging with emerging technologies. Effective AI integration requires calibrated structure that supports critical thinking, along with explicit instruction on evaluating AI outputs against learning objectives as a core digital literacy competency.

Keywords: generative AI, authentic assessment, scaffolding, computing education, student engagement

iafor
The International Academic Forum
www.iafor.org

Introduction

The integration of generative artificial intelligence (AI) within higher education has prompted renewed and continued scrutiny of assessment design, particularly in computing disciplines where both procedural and conceptual understanding are required (Biscalchin et al., 2025). Recent discussions emphasize the need to focus on the underlying assessment construct, ensuring that tasks accurately measure intended learning outcomes rather than peripheral skills or tool usage (Hao et al., 2024). This study presents a comparative case analysis of two approaches to authentic assessment implemented in a second-year module on object-orientation, algorithms, and data structures.

The study examines how differing levels of scaffolding and autonomy influence student engagement, understanding, and performance when generative AI is incorporated into assessment tasks (Prompiengchai et al., 2025). Specifically, it contrasts a structured assessment model supported by a constrained AI chatbot for support with an open-ended model permitting unrestricted AI usage for both requirement generation and implementation.

By analyzing these approaches, the study reiterates the need for more robust, policy-aware approaches to AI integration in assessment design, including human-centered frameworks that ensure alignment between technological tools and pedagogical intent (Biscalchin et al., 2025). This comparison provides insight into how assessment design choices shape student experience and highlights the importance of aligning AI use with clearly defined learning objectives and assessment constructs (Hao et al., 2024; Prompiengchai et al., 2025).

Literature Review

The rapid adoption and accessibility of generative artificial intelligence (AI) tools in education have generated significant debate concerning their pedagogical implications, particularly in relation to assessment integrity, student learning, and assessment design (Biscalchin et al., 2025; Hao et al., 2024). Recent work highlights the need to rethink traditional assessment approaches to ensure alignment with learning outcomes in the presence of generative AI (Hao et al., 2024; Khan et al., 2023).

A central concern is ensuring that assessment tasks remain focused on the intended construct (Hao et al., 2024). In AI-integrated environments, the distinction between student-generated and AI-assisted output can become blurred, raising concerns around validity and authenticity (Hao et al., 2024; Khan et al., 2023). Emerging frameworks emphasize the importance of aligning AI use with pedagogical intent through structured and transparent design principles (Biscalchin et al., 2025; Prompiengchai et al., 2025).

The literature also highlights the importance of scaffolding when students engage with complex tasks and emerging technologies. Without sufficient guidance, students may face increased cognitive load and struggle to manage both the conceptual and operational demands of AI-enhanced tasks (Biscalchin et al., 2025).

Furthermore, authentic assessment in the age of AI must shift focus from product to process. Emphasizing how students arrive at solutions rather than solely evaluating outputs has been identified as a key strategy for maintaining academic integrity and supporting meaningful learning (Khan et al., 2023).

The ability to critically evaluate AI-generated outputs is increasingly recognized as a core digital literacy skill. Students must develop the capacity to assess the quality, relevance, and appropriateness of AI outputs in relation to learning objectives (Hao et al., 2024; Vadaparty et al., 2025).

These considerations underpin the present study, which examines how varying levels of scaffolding and AI autonomy shape student experiences in authentic assessment contexts.

Methodology

This study adopts a comparative case study design, analyzing two iterations of assessment implemented within the same module context.

In the first iteration, students engaged in structured coursework characterized by predefined requirements and the provision of starter code. A custom AI chatbot was integrated to support comprehension and guide task processes (Prompiengchai et al., 2025). Critically, the chatbot was restricted from generating code, thereby functioning as a pedagogical support tool rather than a solution generator.

In the second iteration, a contrasting design was implemented in which generative AI was explicitly incorporated as part of the assessment process. Students were permitted to use AI tools to generate requirement specifications, effectively simulating interactions with virtual clients.

Students were granted full autonomy to determine which requirements to implement and how to demonstrate learning outcomes relating to design principles and patterns. AI-generated code was permitted, subject to mandatory declaration.

This design enables a direct comparison of structured versus open-ended AI integration, facilitating analysis of their respective impacts on engagement, decision-making, conceptual understanding, and the alignment between student outputs and intended learning outcomes.

Results

The structured AI-assisted assessment demonstrated positive outcomes in both student engagement and learning processes. The chatbot effectively supported students in interpreting requirements and establishing workflows, resulting in a reduction in instructor queries. Assessment outcomes remained consistent with previous cohorts, with an observable shift towards higher classifications.

Although there was some indication of limited AI-assisted code generation, the constrained design mitigated the risk of overreliance. The structured environment facilitated clarity and enabled students to focus on applying conceptual knowledge.

In contrast, the open-ended assessment approach resulted in increased student anxiety and a higher frequency of instructor queries. Students reported difficulty in determining appropriate requirements and ineffectively evidencing learning outcomes.

The absence of predefined structure introduced significant decision-making challenges. Students frequently sought validation and expressed uncertainty regarding scope. Furthermore,

many encountered difficulties in relating AI-generated outputs to underlying design principles, resulting in conceptual misalignment with intended learning outcomes.

Discussion

The findings underscore the importance of maintaining alignment between assessment design and intended learning outcomes in AI-integrated environments. The structured approach ensured that AI functioned as a supportive tool without displacing core learning objectives, consistent with recommendations that emphasize principled AI integration in pedagogy (Biscalchin et al., 2025).

By contrast, the open-ended approach shifted significant cognitive and metacognitive demands onto students. Learners were required to assess requirements, evaluate AI outputs, and demonstrate conceptual understanding simultaneously (Hao et al., 2024). This led to increased cognitive load, supporting findings that autonomy without adequate scaffolding can hinder learning outcomes (Biscalchin et al., 2025).

The study also highlights the importance of clearly defining the pedagogical role of AI. Positioning AI as an assistive tool rather than a generative solution appears to promote deeper engagement with learning processes. In contrast, unrestricted AI use may shift focus towards output production, thereby weakening the validity of assessment (Hao et al., 2024).

Additionally, the results reinforce the importance of developing students' critical AI literacy. Students must be able to interrogate and evaluate AI-generated outputs rather than accept them uncritically. This aligns with emerging views that digital and prompt literacy are essential competencies in AI-enabled education (Hao et al., 2024; Prompiengchai et al., 2025).

Overall, the findings suggest that effective AI integration requires careful calibration between structure and autonomy, ensuring that technological affordances support rather than undermine pedagogical goals.

Conclusion

This study provides a comparative analysis of two approaches to integrating generative AI in authentic assessment design demonstrating that the level of scaffolding significantly influences student experience and learning outcomes.

The structured, scaffolded approach supported student engagement, reduced instructional demands, and facilitated consistent demonstration of learning outcomes. In contrast the open-ended approach characterized by high autonomy and unrestricted AI use, resulted in increased anxiety, decision-making challenges, and difficulties in linking outputs to conceptual understanding.

These findings indicate that the effectiveness of generative AI in education is contingent on the manner of its integration. Structured guidance, clearly defined roles for AI, and explicit support for evaluating AI-generated outputs are essential to ensuring meaningful learning.

The study emphasizes the importance of balancing autonomy with scaffolding in the design of AI-integrated assessments particularly in contexts requiring both technical proficiency and conceptual understanding.

Future assessment design must prioritize structured AI integration, emphasize process-oriented evaluation, and explicitly develop students' ability to critically engage with AI as part of their core disciplinary and digital literacy competencies.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that Microsoft Copilot was used in proofreading and refining the language used in the manuscript for accuracy and clarity. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

References

Biscalchin, A., Sarkheyli-Hägele, A., Eriksson, J., & Vogel, B. (2025). Designing Education for the AI Era: Principles for Integrating LLMs in Pedagogy. *2025 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, 1–6. https://doi.org/10.23919/softcom66362.2025.11197458

Hao, J., von Davier, A. A., Yaneva, V., Lottridge, S., von Davier, M., & Harris, D. J. (2024). Transforming Assessment: The Impacts and Implications of Large Language Models and Generative AI. *Educational Measurement*, *43*(2). https://doi.org/10.1111/emip.12602

Khan, M. M., Dong, Y., & Manesh, N. A. (2023). Authentic Assessment Design for Meeting the Challenges of Generative Artificial Intelligence. *IEEE Frontiers in Education Conference (FIE)*. https://doi.org/10.1109/fie58773.2023.10343376

Prompiengchai, S., Narreddy, C., & Joordens, S. (2025). *A Practical Guide for Supporting Formative Assessment and Feedback Using Generative AI*. ArXiv.org. https://arxiv.org/abs/2505.23405

Vadaparty, A., Smith IV, D., Srinath, S., Padala, M., Alvarado, C., Benario, J. G., Zingaro, D., & Porter, L. (2025). *Integrating Large Language Models and Evaluating Student Outcomes in an Introductory Computer Science Course*. ArXiv.org. https://arxiv.org/abs/2510.18806

Contact email: vaidhiyanathanr@cardiff.ac.uk