

Commentary: ChatGPT-supported student assessment – can we rely on it?

The advent of generative AI, particularly ChatGPT, in the educational sector is sparking a multifaceted debate on its reliability for student assessment. Recent literature focuses on the potential of ChatGPT to revolutionize assessment methods, raising both enthusiastic endorsements for its capabilities and serious concerns about the integrity and effectiveness of AI-driven evaluations (Klyshbekova and Abbott, 2024).

ChatGPT's advanced natural language processing (NLP) capabilities can mimic human-like text generation, suggesting a shift from traditional, linear models of learning assessment toward more dynamic and personalized approaches (Damaševičius, 2023). This transition could enhance learning outcomes by tailoring educational experiences to individual needs and providing immediate feedback. The reliability of such AI assessments is under scrutiny, particularly concerning the depth and academic rigor of the responses.

The integration of ChatGPT for student grading has presented various challenges, juxtaposing traditional assessment methods against emerging AI-driven approaches (Jukiewicz, 2024). ChatGPT when used for grading can handle large volumes of assessments quickly, providing immediate feedback that is consistent as long as the input remains within the model's training data scope (Kooli and Yusuf, 2024). ChatGPT can grade assignments across different subjects, demonstrating a moderate correlation with human graders, which highlights its potential as a supportive tool (Kooli and Yusuf, 2024). However, ChatGPT's grading performance can lack depth, missing nuanced insights that experienced educators might offer (Ghapanchi and Purarjomandlangrudi, 2023). The complexity and subtlety of student responses can be underappreciated by AI, which may struggle with interpretations that require deep understanding and contextual awareness. The standardization of feedback may not address specific developmental needs of individual students, which are critical for personalized learning.

The adoption of ChatGPT and similar AI technologies in student grading and assessment heralds significant broader implications for the field of education (Yang *et al.*, 2023). One of the most transformative impacts is the potential shift in the role of educators. With AI handling more routine and administrative tasks, educators can redirect their focus towards more in-depth, interactive and personalized teaching methods. This could lead to an enhancement of the pedagogical process, where the emphasis shifts from teaching to the test to fostering deeper understanding and critical thinking skills (Damaševičius and Sidekierskiene, 2024).

The integration of AI into assessment processes challenges and could reshape the traditional assessment paradigms. AI's capability to analyze vast amounts of data can lead to the development of more sophisticated and adaptive learning environments, where



assessments are customized to the needs of individual students (Alabidi *et al.*, 2023). This could accelerate the move from a one-size-fits-all approach to a more personalized, learner-centered approach in education, promoting greater equity in learning opportunities.

The reliance on AI for educational assessments raises critical ethical questions, particularly regarding fairness and privacy. The potential for AI to introduce biases, whether through its datasets or algorithms, poses a significant risk of perpetuating inequalities (Kooli and Yusuf, 2024). Ensuring that AI systems are fair and transparent is crucial to maintaining trust in educational outcomes.

The broader use of AI in education also demands a reassessment of skills that are taught and valued in educational systems. As AI takes over more routine cognitive tasks, there is a growing need to enhance skills that are uniquely human, such as creative problem-solving, empathy and interpersonal communication (Kirwan, 2023). This shift could lead to significant changes in curriculum design and teaching strategies, emphasizing skills that prepare students for a future where AI is a ubiquitous part of the professional landscape.

One major avenue of future research is the development of enhanced AI models that can better understand and evaluate complex student responses. Current models, while effective in handling straightforward tasks, often struggle with the nuances of higher-order thinking and creativity expressed in student work. Research aimed at improving AI's cognitive and evaluative capabilities could bridge this gap, making AI grading and feedback more comparable to that provided by skilled human educators (Divason *et al.*, 2023).

Another critical area of research involves investigating the long-term effects of AI-assisted education on student learning outcomes. Such research would offer valuable insights into the effectiveness of AI tools in boosting student engagement, retention and achievement, as well as their potential to reduce educational disparities (Alabidi *et al.*, 2023).

Finally, studies on the psychological impacts of AI interactions in educational settings could provide a deeper understanding of how these technologies affect student motivation, trust and perception of learning.

Robertas Damaševičius

Vytautas Magnus University, Kaunas, Lithuania

References

- Alabidi, S., Alarabi, K., Alsalthi, N.R. and Mansoori, M.A. (2023), "The dawn of chatgpt: Transformation in science assessment", *Eurasian Journal of Educational Research*, No. 106, pp. 321-337, doi: [10.14689/ejer.2023.106.019](https://doi.org/10.14689/ejer.2023.106.019).
- Damaševičius, R. (2023), "The rise of chatgpt and the demise of bloom's taxonomy of learning stages", in *Creative AI Tools and Ethical Implications in Teaching and Learning*, pp. 115-134, doi: [10.4018/979-8-3693-0205-7.ch006](https://doi.org/10.4018/979-8-3693-0205-7.ch006).
- Damasevicius, R. and Sidekerskiene, T. (2024), "Ai as a teacher: a new educational dynamic for modern classrooms for personalized learning support", *AI-enhanced Teaching Methods*, pp. 1-24, doi: [10.4018/979-8-3693-2728-9.ch001](https://doi.org/10.4018/979-8-3693-2728-9.ch001).
- Divason, J., de Pison, F.J.M., Romero, A. and de, E.S. (2023), "Cabez 'on. Artificial 'intelligence models for assessing the evaluation process of complex student projects", *IEEE Transactions on Learning Technologies*, Vol. 16 No. 5, pp. 694-707, doi: [10.1109/TLT.2023.3246589](https://doi.org/10.1109/TLT.2023.3246589).
- Ghapanchi, A.H. and Purarjomandlangrudi, A. (2023), "Chatgpt and generative ais: what it means for academic assessments", volume 2023-March.
- Jukiewicz, M. (2024), "The future of grading programming assignments in education: the role of chatgpt in automating the assessment and feedback process", *Thinking Skills and Creativity*, Vol. 52, 101522, doi: [10.1016/j.tsc.2024.101522](https://doi.org/10.1016/j.tsc.2024.101522).

- Kirwan, A. (2023), "Chatgpt and university teaching, learning and assessment: some initial reflections on teaching academic integrity in the age of large language models", *Irish Educational Studies*, pp. 1-18, doi: [10.1080/03323315.2023.2284901](https://doi.org/10.1080/03323315.2023.2284901).
- Klyshbekova, M. and Abbott, P. (2024), "Chatgpt and assessment in higher education: a magic wand or a disruptor?", *Electronic Journal of E-Learning*, Vol. 22 No. 2, pp. 30-45, doi: [10.34190/ejel.21.5.3114](https://doi.org/10.34190/ejel.21.5.3114).
- Kooli, C. and Yusuf, N. (2024), "Transforming educational assessment: insights into the use of chatgpt and large language models in grading", *International Journal of Human-Computer Interaction*, pp. 1-12, doi: [10.1080/10447318.2024.2338330](https://doi.org/10.1080/10447318.2024.2338330).
- Yang, X., Wang, Q. and Lyu, J. (2023), "Assessing chatgpt's educational capabilities and application potential", *ECNU Review of Education*, doi: [10.1177/20965311231210006](https://doi.org/10.1177/20965311231210006).