Weighing it up: Potential benefits and risks of AI assessment

Article #4 of AI in Education Article Series: February 2025

As discussed in our previous article, Artificial Intelligence (AI) can be used to address various challenges with traditional assessment – and in many different ways. Our project is most focused on developing a learner-focused agent – specifically, an agent that is trained on course materials and existing assessment rubrics to run an oral assessment with a learner and grade them. It is important for us to understand both the benefits and risks of this.

This article is the **fourth in a series titled "AI in Education"**, aimed at education providers interested in AI. The intention is for this series to act as a beginner's guide to the use of AI in education, with a particular focus on AI agents. This series is being developed as part of a project to develop an AI agent for learner oral assessment, funded by the Food and Fibre Centre of Vocational Excellence. We invite you to follow along as we (Scarlatti) document our learnings about this exciting space.

The article below provides an overview of the **potential benefits and risks of learner-focused agents.** We suggest these *speculatively*, rather than with the intention of being definitive or exhaustive. This is because it is based on industry stakeholder conversations and early desk research, rather than pilots (scheduled for April 2025). We believe the real benefits and risks will become clearer after these pilots.

Potential benefits

We suggest that an AI agent for oral assessment may provide benefits, such as:

- **Greater personalisation -** In the same way that a human assessor can adapt their approach to suit the learner during an assessment, an AI agent could adapt as the assessment proceeds. This offers the potential for 'mass-personalisation' of the assessment process.
- **Real-time feedback and coaching** – Relating to the above, the conversational nature of AI agents means that they could be capable of adapting to learners' questions and providing

personalised feedback. This could help students who are not comfortable asking for support, or who cannot afford an additional tutor.

- **Improved inclusivity –** Learners will provide their answers verbally to the agent. This may benefit learners who have learning difficulties, neurodivergence or English as a second language.
- **Increased enrolments -** People who are deterred from formal study due to the need for written assessments may be more interested in applying to courses that harness these AI agents.
- **Improved flexibility** – As these agents could be accessed remotely (via the internet), learners could have more flexibility to choose when, where and how they carry out assessments. For instance, a student could talk to the agent to undertake an assessment while milking.
- **Increased digital literacy –** Using this technology could aid learners in developing greater digital literacy skills, including familiarity with AI tools.

Potential risks

However, adopting AI assessment agents may also pose risks, such as:

- **Privacy concerns** – Some AI models use the data that users feed into the agent to continue to train the model, which is concerning for learners' inputting sensitive or personal information. This is a risk that can be managed using different premium subscriptions to the AI model.
- **Inability for the agent to understand users' accents** – There is also a risk when AI agents are conducting verbal assessments that they could struggle to understand different learners' accents, leading to issues determining whether they have answered the question correctly and deciding on preliminary grades.
- **Inaccurate results –** Depending on how the agent is designed and what content is inputted into it (i.e. assessment answers, marking rubrics, course content) the agent could misinterpret what the learner is trying to say which could lead to unreliable results. This is particularly concerning for assessments that are set up to only pass learners who mention specific terms in their answers.
- **Bias in AI models**[1] – AI agents could perpetuate biases in the AI model's training data and could lead them to make discriminatory decisions during the assessment. This is particularly relevant in the context of New Zealand as an AI model originating in the United States is unlikely to have a comprehensive understanding of Mātauranga Māori or Te Reo.
- **Data exposure**[2] – As with any technology, there are also concerns that AI agents could be hacked during assessments or disclose learners' information without their consent.

[1] We suggest this risk is secondary to the previous ones in this list, given that modern AI models have bias deliberately removed. We are currently unaware of any work that demonstrates clear identity-based biases that would be material for our use-case.
[2] We suggest this risk is secondary to the previous ones in this list, given that data exposure risk is no worse than any other cloud-based service.

Scarlatti's take

We believe, depending on the problem at hand, that the potential benefits are enough to warrant at least the *exploration* of using an AI assessment agent for learners. Within our project, we will explore current guidance, build in mitigations to risks, and collect feedback from different stakeholders.

Questions that we are asking for our own AI agent:

- Where can we look for guidance on risks? (e.g., government, thought leaders).
- How can we mitigate these risks within our development and pilots?
- How will these benefits and risks be realised during pilots and will the benefits be enough?

Interested in following our journey into AI?

- Sign up to receive our next article directly to your inbox.
- Contact the Scarlatti team to share your thoughts or questions.

Further reading

AI in Higher Education. (2024). *20204 AIHE NZ Symposium Morning Session Summaries.* Retrieved November 26, 2024, from https://aihe.nz/wp-content/uploads/2024/09/Morning-session-summaries-and-data.pdf

Allen, J. (2024, November 25). *The future of apprenticeships: How AI is changing the game.* Retrieved November 26, 2024, from https://www.graide.co.uk/blog/the-future-of-apprenticeships-how-ai-is-changing-the-game

Department of Internal Affairs, National Cyber Security Centre & Stats NZ. (2023). Initial Advice on Generative Artificial Intelligence in the public service. Retrieved January 8, 2025, from https://www.digital.govt.nz/assets/Standards-guidance/Technology-and-architecture/Generative-AI/Joint-System-Leads-tactical-guidance-on-public-service-use-of-GenAI-September-2023.pdf

Kuhail, M.A., Alturki, N., Alramlawi, S., & Alherjori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies* 28, 973–1018. https://doi.org/10.1007/s10639-022-11177-3

Ministry of Education. (2024, November 25). *Generative AI: Guidance and resources for education professionals on the use of artificial intelligence in schools.* Retrieved January 8, 2025, from https://www.education.govt.nz/school/digital-technology/generative-ai

Mollick, E., & Mollick, L. (2023). *Assigning AI: Seven Approaches for Students, with Prompts.* ResearchGate. https://www.researchgate.net/publication/371728434_Assigning_AI_Seven_Approaches_for_Students_with_Prompts

Office of the Privacy Commissioner. (2023). Artificial intelligence and the information privacy principles. Retrieved January 8, 2025, from https://privacy.org.nz/assets/New-order/Resources-/Publications/Guidance-resources/AI-Guidance-Resources-/AI-and-the-Information-Privacy-Principles.pdf

Sáncez-Prieto, C. J., Gamazo, A., Cru-Benito, J., Therón, R., & García-Peñalvo, J. F. (2020). AI-Driven assessment of students: Current uses and research trends. Springer Nature. https://doi.org/10.1007/978-3-030-50513-4_22

Williams, J., & Ang, C. (2024). *Taking Stock: Defining an Excellence Framework for Vocational Education.* Food and Fibre Centre of Vocational Excellence. Retrieved November 26, 2024, from https://foodandfibrecove.nz/wp-content/uploads/2022/02/FFCoVE-Taking-Stock-Lit-review-v6_6-FINAL.pdf