AI and student assessment in human-centered education

Bruno Balducci*

Learning and Teaching, Otago Polytechnic Auckland International Campus, Auckland, New Zealand

1 Introduction

Sustainable Development Goal #4 is about providing quality education and learning opportunities for all. "Quality" is also indicated in Target 4.3 with reference to technical, vocational and tertiary education, while educational content is described in Target 4.4 as skills development "for employment, decent jobs and entrepreneurship", and in Indicator 4.7.1 as "global citizenship education" and "education for sustainable development" (UNESCO, 2014). Underlying all these targets is the ability to achieve learning outcomes and measure progress through valid and reliable student assessment.

This ability has now come under threat from the rapid advances made by generative artificial intelligence (AI), i.e., "technology that (i) leverages deep learning models to (ii) generate human-like content (e.g., images, words) in response to (iii) complex and varied prompts" (Lim et al., 2023, p. 2). AI tools such as ChatGPT are seen to "pose a significant threat to the academic integrity of traditional assessments" (Rudolph et al., 2023a, p. 365–366)—especially, but not exclusively, in higher education. There is deep skepticism regarding AI detection tools (Perkins, 2023) and closed-book exams as an alternative assessment strategy (Rudolph et al., 2023b). Overall, as generative AI continues to evolve, the security of many different types of assessments involving text, images, calculations and even computer code will be increasingly at risk, thus undermining the effectiveness of assessment processes, the usefulness of achievement rates (and associated data) and the motivation of students to learn.

The aim of this opinion piece is to propose the application of humanistic principles of task design as a protective measure against the incursion of AI-generated material in the assessment of learning. As noted by Webb et al. (2022, p. 129), the term "human-centered learning" is being used in the literature on AI to denote "pedagogies that are personally relevant to learners and encourage their active learning and engagement". But although this article addresses the security of assessment tasks, the design of learning tasks should also be considered since "students need the opportunity to practice with the form of assessment before it is used as an assessment" (Gulikers et al., 2004, p. 71). By concentrating on the use of skills and knowledge by real people in realistic situations, as opposed to academic performance in artificial circumstances, student assessment can be safeguarded in the future. But before humanistic strategies are deployed, the principles on which they rest have to be set out and critically examined, an objective this article will seek to fulfill with clarity and concision. A more detailed investigation of how such principles could be implemented in particular disciplines, meanwhile, inevitably remains beyond its scope.

2 Principles of task design

In the face of the new challenges presented by AI, assessments that require creativity or critical thinking are often cited as possible solutions (Ifelebuegu, 2023; Kasneci et al., 2023; Rudolph et al., 2023b). In fact, these capabilities were seen as desirable graduate outcomes long before the arrival of ChatGPT. They were valued on a par with professional,

problem-solving, collaborative and communication skills by Biggs and Tang (2011), who took pains nevertheless to emphasize they should be "embedded" in relevant, discipline-specific contexts. Their advice is now doubly important in that AI, while it lacks conceptual understanding, is capable of generating material that may easily appear creative and innovative, or critical and reflective. In other words, it is contextualization that is needed first and foremost, whereas a focus on creative and critical thinking will not in itself guarantee the safety of assessments.

The question then is: What kind of context is most appropriate for student assessment in the age of AI? And the answer, of course, will depend on the discipline itself. Different subject areas require different methods and approaches. Yet the guiding principle for task designers, regardless of the program involved, must be that it is authentic, thereby avoiding the artificiality of conventional assessments. Ifelebuegu (2023) describes authentic assessment design as "contextual" and "problem-based" (p. 4). Rudolph et al. (2023b) point out that authentic assessment is set "in realistic situations" (p. 355). Gulikers et al. (2004) go further and explain that it "confronts students with activities that are also carried out in professional practice" (p. 71), and should even resemble "social processes that are present in real-life contexts" (p. 74). This will be as much the case for a "theoretical" discipline like mathematics as it is for applied business management. After all, mathematics is used in a wide range of professions and real-world settings which can provide the basis for developing and assessing mathematical ability. Through the careful alignment with learning objectives, a task designer can ensure that assessments are always valid and entirely focused on the relevant discipline.

The social aspect of assessment, as mentioned by Gulikers et al. (2004), brings another key principle to bear, namely collaboration. The ability to work collaboratively with others through teamwork and personal interaction is a skill that transcends knowledge gained from discipline-based education. It is, besides, a type of graduate outcome that education providers have been seeking to address (Biggs and Tang, 2011). Some commentators have identified collaborative assessment as a means of protecting academic integrity from the threat of AI (Ifelebuegu, 2023; Rasul et al., 2023; Rudolph et al., 2023a). Their main justification for adopting collaborative forms of assessment is found in "social constructivist theory's emphasis on collaboration and social interaction in learning" (Rasul et al., 2023, p. 46). As a point of view, it is certainly consonant with the call for assessments which incorporate social processes. Even so, we may question the authenticity of the constructivist model of learning where, in addition to "social collaboration with others", knowledge is constructed "through interactions with teachers" (Rasul et al., 2023, p. 47). This can hardly be represented as mirroring the social relationships that students will find outside of education.

The authenticity of the learning process and its repercussions on modes of assessment enables a third principle to come into play, one that relates to the complexity of real-life problems. Although these may at times require solutions that are simple and unique, they are for the most part complicated and "ill-structured", allowing for a multiplicity of possible solutions (Herrington and Herrington, 1998). Students therefore need to apply higher-order thinking skills to deal with this complexity (Ifelebuegu, 2023), and to have "ownership of the task and the process to develop

a solution" (Gulikers et al., 2004, p. 71). This in turn affects the way that student work is assessed, as well as how the role of AI is perceived. For as Ifelebuegu (2023) remarks, the emphasis now "is on process as much as product" (p. 4), i.e., the work is assessed not only by its product(s), but through the process whereby this outcome is achieved. Although AI might still be involved (as will become increasingly the case in future), the focus remains firmly on the use that is made of it, as opposed to what it generates. The processes that emerge from the complexity of authentic tasks, whether collaborative or not, create opportunities for tracking or overseeing work accomplished by students and preventing the misuse of chatbot technology.

3 Human-centered education in practice

We have, therefore, three principles of task design— authenticity, collaboration, and process—on which the security of assessment is most likely to depend. There may be others that will play a part, and of course the design principles themselves can be formulated or interpreted from different perspectives. The basic strategy, however, remains the same, which is to devise assessable activities around things that humans do and human ways of doing them, instead of merely productive tasks without real world equivalents that may be done just as well by generative AI.

What such assessments mean in practice will vary according to educational aims and contexts, local circumstances and how the factors outlined above are interconnected. As we have seen, an assessment task need not always be collaborative, and what is more, there are diverse forms of cooperation that can be included in its design. Beyond the process/outcome distinction, where the final product is much less significant in the first than in the second, the objective might be to "pinpoint aspects of an individual's behaviors within a group task" (Child and Shaw, 2016, p. 19). Hence collaborative assessment can provide more than a means of "rewarding students for their collective effort, negotiation skills, and ability to reach consensus" (Ifelebuegu, 2023, p. 4). It is possible, in actual fact, to observe a collaborative process for the purposes of assessing individual performance within a realistic setting.

The shape, content, and scope of innovative, human-centric assessment are all adaptable to multiple kinds of design. Existing student-centered assignments might provide models for new tasks that could be tailored to accommodate the changing reality of AI. It is thought that the practice of project management, for example, will be highly impacted by AI technology. But projects can still be used as a template for assessments that prioritise human skills such as leadership and team management (Fridgeirsson et al., 2021). Similar conclusions can be drawn about data science (Tu et al., 2023).

4 Conclusion

To ensure that quality education is maintained despite the ever-greater challenges that AI will undoubtedly bring about, teachers and educators need (i) authentic, (ii) collaborative, and

(iii) process-driven assessments of student learning where the true capacity of people to apply their skills and knowledge in real-world circumstances may be recognised. This article has outlined key humanistic principles for the future security of student assessment, as well as important factors to bear in mind in the corresponding task designs:

- evidence of creative and critical thinking can be imitated by AI
- students' interaction with their peers takes precedence over interaction with teachers
- individual student performance can be assessed through close monitoring of collaborative processes.

Staying focused on real human interaction must be at the heart of a successful integration of AI in the education of tomorrow. In that brave new world, the ability to separate human from artificial in student assessment will become imperative.

Author contributions

BB: Conceptualization, Investigation, Methodology, Writing—original draft, Writing—review & editing.

Conflict of interest

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

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.

References

Biggs, J., and Tang, C. (2011). *Teaching for Quality Learning at University*. New York, NY: McGrawHill

Child, S., and Shaw, S. (2016). Collaboration in the 21st century: implications for assessment. *Res. Matters* 22, 17–22. Available online at: https://api.semanticscholar.org/CorpusID:218883404

Fridgeirsson, T. V., Ingason, H. T., Jonasson, H. I., and Jonsdottir, H. (2021). An authoritative study on the near future effect of artificial intelligence on project management knowledge areas. *Sustainability* 13:4. doi: 10.3390/su13042345

Gulikers, J., Bastiaens, T., and Kirschner, P. (2004). A five-dimensional framework for authentic assessment. *Educ. Technol. Res. Dev.* 52:3. doi: 10.1007/BF02504676

Herrington, J., and Herrington, A. (1998). Authentic assessment and multimedia: How university students respond to a model of authentic assessment. *High. Educ. Res. Dev.* 17, 305–322. doi: 10.1080/0729436980170304

Ifelebuegu, A. (2023). Rethinking online assessment strategies: authenticity versus AI chatbot intervention. *J. Appl. Learn. Teach.* 6:2. doi: 10.37074/jalt.2023.6.2.2

Kasneci, E., Seßler, K., Kuechemann, S., Bannert, M., Dementieva, D., Fischer, F., et al. (2023). *ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education*. Charlottesville, VA: Center for Open Science, EdArXiv.

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., and Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *Int. J. Manag. Educ.* 21:2. doi: 10.1016/j.ijme.2023.100790

Perkins, M. (2023). Academic integrity considerations of AI large language models in the post pandemic era: ChatGPT and beyond. *J. Univ. Teach. Learn. Pract.* 20:2. doi: 10.53761/1.20.02.07

Rasul, T., Nair, S., Kalendra, D., Robin, M., de Oliveira Santini, F., Ladeira, W., et al. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *J. Appl. Learn. Teach.* 6:1. doi: 10.37074/jalt.2023.6.1.29

Rudolph, J., Tan, S., and Tan, S. (2023a). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond: the new AI gold rush and its impact on higher education. *J. Appl. Learn. Teach.* 6:1. doi: 10.37074/jalt.2023.6.1.23

Rudolph, J., Tan, S., and Tan, S. (2023b). ChatGPT: bullshit spewer or the end of traditional assessments in higher education? *J. Appl. Learn. Teach.* 6:1. doi: 10.37074/jalt.2023.6.1.9

Tu, X., Zou, J., Su, W., and Zhang, L. (2023). What should data science education do with large language models? *ArXiv*. doi: 10.48550/arXiv.2307.02792

UNESCO (2014). *Roadmap for Implementing the Global Action Programme on Education for Sustainable Development*. Paris: United Nations Educational Scientific Cultural Organization. Available online at: https://unesdoc.unesco.org/ark:/48223/pf0000230514 (accessed November 9, 2023).

Webb, S., Hodge, S., Holford, J., Milana, M., and Waller, R. (2022). Aligning skills and lifelong learning for human-centred sustainable development. *Int. J. Lifelong Educ.* 2, 127–132. doi: 10.1080/02601370.2022.2057167