11th International Conference on Higher Education Advances (HEAd'25)
Universitat Politècnica de València, València, 2025
DOI: https://doi.org/10.4995/HEAd25.2025.19985

Assessments - search for an answer in the era of Generative AI

Ali Hellany[1], Anush Heinitz[1], Iain McAlpine[2], Karen Yevenes[1], Matt Stevens[1], Surendra Shrestha[1]

[1]School of Engineering, Design and Built Environment, Western Sydney University, Australia. [2]Office of the Pro Vice-Chancellor, Quality and Integrity, Western Sydney University, Australia.

How to cite: Hellany, A.; Heinitz, A.; McAlpine, I.; Yevenes, K.; Stevens, M.; Shrestha, S. (2025). Assessments - search for an answer in the era of Generative AI. In: 11th International Conference on Higher Education Advances (HEAd'25). Valencia, 17-20 June 2025. https://doi.org/10.4995/HEAd25.2025.19985

Abstract

Assessing student learning has always been challenging. The recent proliferation of generative artificial intelligence (GenAI) tools has increased this complexity driving educators to search for innovative and meaningful assessment practices. Educational institutions are investigating, developing and implementing alternative and authentic assessment practices that maintain academic integrity while genuinely assessing student learning. This paper summarises the assessment practices in a techno-centric school at an Australian university and proposes assessment strategies to control academic misconduct cases arising because of GenAI. A teaching and learning case study of authentic assessment to engage students and effectively achieve learning outcomes is discussed. The technique has been popular among the students while successfully navigating the challenges provided by recent developments in GenAI.

Keywords: *academic integrity, authentic assessment, generative AI, GenAI.*

1. Introduction

Assessment is one of the methods used to test student learning and attainment of knowledge. Therefore, it is important to design assessments carefully to encourage student engagement while enhancing student learning (Biggs et al., 2022). Assessment design gained increased importance since the introduction of online assessments forced upon the academic community by the COVID-19 pandemic, when online education underwent a monumental shift (Daniel, 2020). Then came the release of an artificial intelligence tool, ChatGPT (Chat Generative Pre-trained Transformer) in November 2022 (OpenAI, 2022). With these changes came renewed concerns about the integrity of traditional assessments (Yusuf et al., 2024).

The abovementioned factors and the unprecedented increase in academic misconduct cases due to GenAI use were the catalysts that led to the formation of the Alternative Assessment Working

Party (AAWP) in the School of Engineering, Design and Built Environment (SoEDBE) at Western Sydney University (WSU). Established in June 2022, the group was tasked with developing strategies to address academic integrity within the SoEDBE.

According to the Cambridge University Dictionary, the abbreviated term for 'generative artificial intelligence' is 'Generative AI' or 'GenAI'. In this paper, the authors use 'GenAI' to define: *the use of artificial intelligence (= computer systems that have some of the qualities that the human brain has, such as the ability to interpret language, recognize images, and learn from data supplied to them) capable of producing texts, images, etc*. Furthermore, this paper presents initiatives undertaken by the AAWP to identify opportunities for promoting authentic assessment to drive a positive student learning outcome.

2. Literature Review

This section briefly reviews the literature on the role of assessment in student learning and the impact of GenAI on assessment design and delivery.

2.1. Assessment and its role in student learning

Assessments have been identified as a key driver of student learning activities. Biggs et al. (2022) assert that assessment requirements effectively define the curriculum for most students and that assessment requirements are a starting point for student learning activities, more so than teaching activities and subject content. While some academics may consider this as a negative, a focus on assessment can lead to active and effective learning if the assessment tasks are designed effectively (Biggs et al., 2022).

Assessment tasks can be designed to guide students towards analytical thinking and problem-solving, thus leading to higher-order thinking and enabling high-level subject learning outcomes. Boud and Dochy (2010) argue assessment tasks achieving this represent '*assessment for learning*' and should be at the heart of subject and course design so that the effects of assessment are systematically planned and scaffolded. This may involve including industry perspectives in the design of assessment tasks to develop skills important for employability.

A further development is the idea of authentic assessment. According to Quinlan, Sellei & Fiorucci (2024), authentic assessments promote meaningful engagement by students in a range of areas such as professional, societal, disciplinary and developmental contexts. They argue that this helps students see the relevance of subject matter as applied to the real-world professional environment.

Many researchers question the notion of authentic assessment. What are the attributes of an authentic assessment? Which factors need to be considered to make an assessment authentic? (Ajjawi et al., 2024; Messier, 2022; Quinlan et al., 2024). However, there is a general agreement

that authentic assessments engage students and help achieve better learning outcomes (Villarroel et al., 2018). Designed properly using sound educational pedagogy, authentic assessments prepare students to find solutions to real-world problems. The approach can also be compared to proponents who use Conceive, Design, Implement, Operate (CDIO) as a framework for student learning (CDIO, n.d.).

2.2. Assessment and Generative Artificial Intelligence (GenAI)

The public release of ChatGPT by OpenAI (OpenAI, 2022) in November 2022, it has disrupted the way problems are investigated, including in the higher education sector (McDonald et al., 2024; Perkins et al., 2024). The challenges to academia due to such technological advancement and the rapidly expanding commercial availability of GenAI tools have been highlighted by Yesmin (2024). In addition, the human-like responses of AI-powered chatbots and their ability to articulate coherent essays have been of concern in academic spheres (Shahriar & Hayawi, 2024; Yusuf et al., 2024).

GenAI is leading to a significant shakeup in assessment. As Nikolic et al. (2023) identified in a case study, most assessment tasks in engineering-related programs (at one university) could be accomplished to a passing standard and, in some cases, a high standard by using a range of GenAI programs. While programs exist to detect GenAI use, savvy students can re-prompt AI or employ more than one GenAI system so that the GenAI content is undetectable (Liu & Chilton, 2022).

These factors assess learning greater importance and urgency in the higher education environment. In a recent report on the challenges of GenAI (TEQSA, 2024), the authors indicate that assessment tasks designed based on principles that promote effective learning also provide a more robust resistance to the threats posed to academic integrity by GenAI. Assessment tasks engaging students in an extensive process of investigation of sources, analysis, problem-solving, working in partnership with academic staff and other students are less likely to be vulnerable to work produced by GenAI (Lodge et al., 2023).

A focus on the process of learning, rather than an artefact as a product that may or may not be the student's own work, may be a useful assessment approach. In addition, assessments that promote group-based projects with regular reporting, team meetings and journaling to enable self-analysis and reflection of the stages of learning (CSU, n.d.) could enhance assessment planning. Peer review and feedback on group tasks can also enable consideration of the learning process. Authentic tasks involving external partners, such as industry collaborative projects, can also be valuable in enabling consideration of the learning process independently of the final product. Ajjawi et al. (2024) suggest authentic assessment is valuable in addressing three assessment challenges in higher education: the rise of artificial intelligence, threats to academic integrity, and greater student equity.

There is a strong similarity between authentic assessment and alternative assessment (definition provided in section 4 below) as defined by the School's AAWP. The working party aimed to promote alternative assessments and identify the extent to which they are currently used within the School.

3. Methodology

A mixed-methods approach was adopted in this study. Firstly, a quantitative online survey was distributed to all academics in the School in October 2022 to take stock of the types of assessments used in subjects offered by the School. Questions ranged from the number and types of assessments in each subject to the percentage distribution of tasks and procedures students used for task submission. The quantitative survey was followed by written feedback from participants of two workshops to ascertain academic concerns about GenAI and opportunities to redesign assessments. Finally, in the teaching exemplar, student testimonials were gathered in response to one neutral question posed after completing the subject requirements, asking students to write a statement (85 words or less) about their experience working with an industry partner.

4. Survey data collection, response analysis and results

To understand the range of assessment techniques used in the School, the AAWP issued surveys to all academics in SoEDBE who were requested to respond within 2-weeks. Responses were categorized according to assessment types that were '*alternative assessment*' or other types, to provide a snapshot of the proportion of subjects employing the broad definition of alternative assessment approaches.

3.1. Survey response

Responses covering 201 subjects from every discipline in the School were received. The survey addressed a wide range of questions about assessment practice across disciplines. Of particular significance were the range of assessment types in practice in the School (see Figure 1).

The overall assessment pattern heavily relies on traditional assessment types like quizzes, reports, and exams. Alternative formats, such as applied projects and case studies, show possibilities for creating a more widespread use of authentic assessment across the school. The data indicate the extent of the challenge facing the School, given the growing capability of GenAI products to produce effective and mostly undetectable responses to most traditional assessment formats.

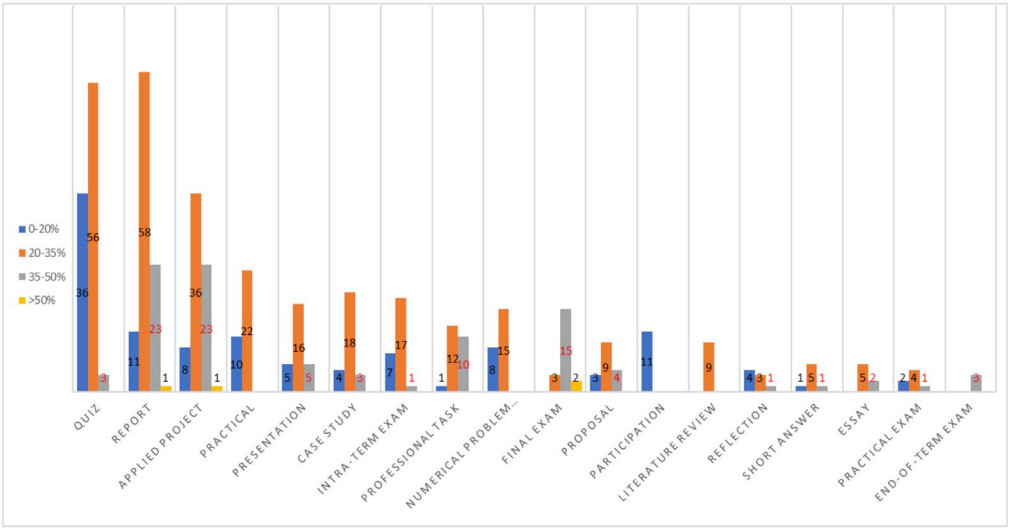

Figure 1. Assessment Types (Traditional vs Alternative)

5. Alternative Assessment Workshops

Following the survey and staff development workshops, a presentation followed whereby small-group discussions were held to promote the idea of alternative assessment and to gauge the response by academic staff to the significance of alternative assessment formats in the current environment. An example is presented below.

5.1. Example – an undergraduate subject using an alternate assessment strategy

In a second-year Industrial Design (ID) subject delivered within the School, assessments were designed to deter exclusive use of GenAI in writing and research. Students were challenged with authentic learning tasks to promote critical thinking and problem-solving; tasks were designed to mimic typical design problems when working in industry. The learning environment (classroom set-up) encompassed small teams rotating to various '*learning stations*' whilst undertaking written, research, and practical tasks. For example, during teamwork scenarios, students participated in mind-mapping at pinboards, project planning using laptops, sketching on tablets, paper or whiteboards, and verbal group presentations; tutors were also physically active, moving between groups and providing guidance. All tasks required students to rely on the creative input of team members, whereby solutions are informed by tangible experiences rather than relying on GenAI.

The subject's content was lively and visible (*Figure 2*); students witnessed concepts come to life and could visualize the application of content to their profession. For example, one of the

learning objectives of the subject was the application of anthropometric data to inform the design of an object. This was achieved through students' use of biometric measuring tools to capture data from peers, a task that cannot be repeated using GenAI. Students also learnt principles of ethics in engaging with participants in addition to presenting and analyzing the data using appropriate statistical techniques. As a result, the content was immediately placed in context of the students' world without relying on GenAI software.

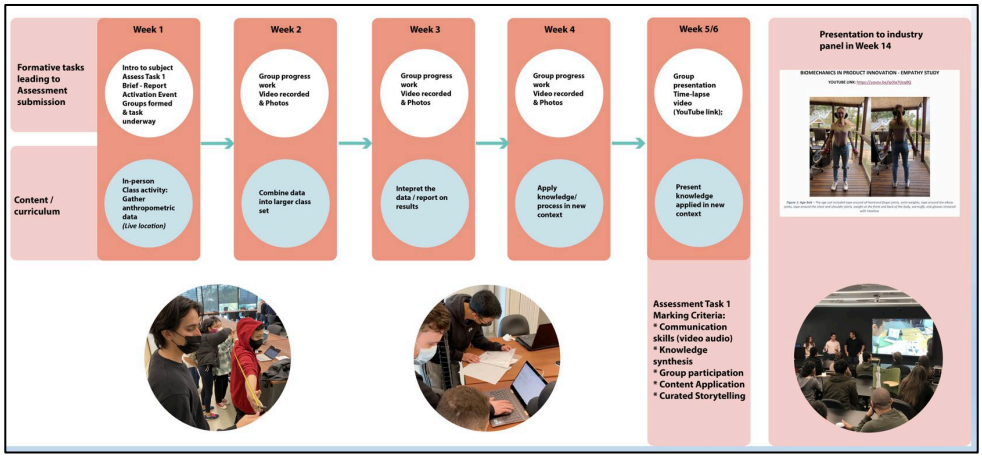

Figure 2. Learning and teaching journey in an Industrial Design subject

Student participation was self-recorded. They kept audio-visual recordings (or A/V journals) to capture processes of their own participation in all activities. The footage was curated and compiled as a time-lapsed video (using basic video-editing software) and shared with the industry partner and academics. Audio-visual journaling provides evidence of each student's contribution to their group's design solution. Students know they must show their workings in the footage, so GenAI avatars are difficult to replicate in team meetings.

An industry partner collaborated with students and co-designed the project to ensure the subject unfolded in a manner mimicking a real-world scenario. During the problem-solving and design conceptualization phases, the industry partner provided professional perspectives and critiqued students' progress. This engagement was invaluable in driving students to enhance their creative approaches to the project. In addition, the industry partner facilitated a field trip for students, providing an opportunity for students to observe design processes and robotic manufacturing in an industry setting.

The student feedback was highly positive, suggesting student learning through interactive classroom experiences, varied assessment tasks, and industry partnership. Selected comments from two students on aspects of the subject they liked or did not like are provided below:

" *[the industry partnership] was a rich learning experience. Collaborating with our industry partner provided valuable insights into real-world applications. I particularly appreciated the emphasis on precision and safety [during the field trip], ensuring the devices met high standards. The* interdisciplinary *nature of the project enhanced my problem-solving skills, making it a beneficial endeavor in understanding (subject name) principles within a practical context.*" (student A)

"*Engaging in a diverse range of (subject name) activities this semester, such as an industry project, an empathic* study *and some quizzes, has been amazing. Additionally, the collaboration with an industry partner has provided relevant and practical insights, allowing an alignment of theoretical concepts with real-world applications. Furthermore, the experiential learning, from anthropometrics data gathering to industry presentations, improved my understanding of this subject's broader implications."* (student B)

The above comments suggest that students gained a rich and deep study experience and that the learning objectives were achieved authentically without the use of GenAI.

6. Conclusions and recommendations

The proliferation of GenAI tools and their widespread unethical use in the higher education sector has given rise to academic integrity issues. While this has challenged educators, they have risen to the occasion by developing innovative learning and teaching techniques and novel assessment practices. Higher education institutions and regulatory agencies have led by example through development and implementation of new guidelines and policies to ensure that the students are prepared for the profession they will be serving after graduation. The successful strategies described and student feedback in one of the subjects mentioned in this paper are testimony to the academics' innovation capability. Many more appropriate strategies must be developed and implemented in other subjects in different courses. These can then be incorporated to improve curricula while benefiting from GenAI.

References

Ajjawi, R., Tai, J., Dollinger, M., Dawson, P., Boud, D., & Bearman, M. (2024). From authentic assessment to authenticity in assessment: broadening perspectives. *Assessment & Evaluation in Higher Education*, *49*(4), 499-510.

Biggs, J., Tang, C., & Kennedy, G. (2022). *Teaching for quality learning at university 5e*. McGraw-hill education (UK).

Boud, D., & Dochy, F. (2010). Assessment 2020. Seven propositions for assessment reform in higher education.

CDIO. (n.d.). *Conceive, Design, Implement and Operate*. Retrieved 23 January 2025 from https://cdio.org/

CSU. (n.d.). *Rethinking Assessment Strategies in the Age of Artificial Intelligence (AI)*. Division of Learning and Teaching. Retrieved 23 January 2025 from

https://www.csu.edu.au/division/learning-teaching/assessments/assessment-and-artificial-intelligence/rethinking-assessments?t

Daniel, S. J. (2020). Education and the COVID-19 pandemic. *PROSPECTS*, *49*(1), 91-96. https://doi.org/10.1007/s11125-020-09464-3

Liu, V., & Chilton, L. B. (2022). *Design Guidelines for Prompt Engineering Text-to-Image Generative Models* Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, New Orleans, LA, USA. https://doi.org/10.1145/3491102.3501825

Lodge, J., Howard, S., & Bearman, M. (2023). Assessment reform for the age of artificial intelligence.

McDonald, N., Johri, A., Ali, A., & Hingle, A. (2024). Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. *arXiv preprint arXiv:2402.01659*.

Messier, N. (2022). *Authentic Assessments*. Center for the Advancement of Teaching Excellence at the University of Illinois Chicago. Retrieved 23 January 2025 from https://teaching.uic.edu/resources/teaching-guides/assessment-grading-practices/authentic-assessments/

Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G. M., Grundy, S., Lyden, S., Neal, P., & Sandison, C. (2023). ChatGPT versus engineering education assessment: a multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity. *European Journal of Engineering Education*, *48*(4), 559-614. https://doi.org/10.1080/03043797.2023.2213169

OpenAI. (2022). *Introducing ChatGPT*. Retrieved 23 January 2025 from https://openai.com/index/chatgpt/

Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The Artificial Intelligence Assessment Scale (AIAS): A Framework for Ethical Integration of Generative AI in Educational Assessment. *Journal of University Teaching & Learning Practice*, *21*(6), 1-18. https://doi.org/10.53761/q3azde36

Quinlan, K. M., Sellei, G., & Fiorucci, W. (2024). Educationally authentic assessment: reframing authentic assessment in relation to students' meaningful engagement. *Teaching in Higher Education*, 1-18.

Shahriar, S., & Hayawi, K. (2024). Let's have a chat! A Conversation with ChatGPT: Technology, Applications, and Limitations. Artificial Intelligence and Applications,

TEQSA. (2024). *Gen AI strategies for Australian higher education: Emerging practice*. Tertiary Education Quality and Standards Agency. Retrieved 25 January 2025 from https://www.teqsa.gov.au/sites/default/files/2024-11/Gen-AI-strategies-emerging-practice-toolkit.pdf

Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: creating a blueprint for course design. *Assessment & Evaluation in Higher Education*, *43*(5), 840-854.

Yesmin, S. (2024). Redesigning tertiary educational evaluation with AI: A task-based analysis of LIS students' assessment on written tests and utilizing ChatGPT at NSTU. *Science & Technology Libraries*, *43*(4), 355-367.

Yusuf, A., Pervin, N., & Román-González, M. (2024). Generative AI and the future of higher education: a threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education*, *21*(1).