CENTRE FOR RESEARCH IN ASSESSMENT AND DIGITAL LEARNING

DEAKIN UNIVERSITY

CRADLE suggests...

Issue 09 | 2023

Assessment and genAI

Assessment in a world of artificial intelligence

The introduction of readily available AI resources provides a catalyst for better assessment practices. Enacting the principles of good assessment becomes even more pressing as we prepare our students to work and live with artificial intelligence.

What is genAI and why does it matter for assessment?

ChatGPT is a prominent example of generative Artificial Intelligence (genAI). It generates coherent and complex responses, based on statistical recognition of existing textual patterns in a large corpus of sources. GenAIs can also produce images, numeric data and references, based on similar types of predictive algorithms. GenAI challenges many assessment designs, particularly those relying on students' knowledge recall and simple analysis, as these were previously considered to be only achievable by humans. Detection of genAI generated outputs has not yet proved reliable.

The research

CRADLE publications outline many aspects of assessment relevant to the introduction of genAI (1,2,3,4,5). We identify two sets of recommendations that enhance student development as well as promoting academic integrity. We make suggestions for: 1) how the principles of assessment design can be enacted in a time of genAI; and 2) how assessment practices can immediately take account of genAI, including ChatGPT.

CRADLE suggests is a series of briefings from the **Centre for Research and Assessment in Digital Learning (CRADLE)**, which translates our own research into practice-based possibilities.

Assessment and genAI continues on page 2

1. Ajjawi, R., Tai, J., Boud, D., & Jorre de St Jorre, T. (Eds) (2023). *Assessment for Inclusion in Higher Education: Promoting Equity and Social Justice in Assessment.* Routledge.
2. Bearman, M., Dawson, P., Ajjawi, R., Tai, J. & Boud, D. (Eds.) (2020) *Re-imagining university assessment in a digital world* Springer.
3. Boud, D., Ajjawi, R., Dawson, P., & Tai, J. (Eds.). (2018). *Developing evaluative judgement in higher education: Assessment for knowing and producing quality work.* Routledge.
4. Dawson, P. (2020). *Defending assessment security in a digital world: preventing e-cheating and supporting academic integrity in higher education.* Routledge.
5. Henderson, M., Ajjawi, R., Boud, D., & Molloy, E. (Eds.). (2019). *The Impact of Feedback in Higher Education: Improving assessment outcomes for learners.* Palgrave Macmillan

Enacting principles of good assessment design, in a world of genAI

Deploy resources to assure assessment when it matters most

Assessment at the most basic level requires us to assure that all students have achieved all learning outcomes at least at a threshold level. This function, as distinct from other purposes of assessment, needs to be protected to a high level of integrity. Assessment for assurance may therefore involve a focus on meeting *program* learning outcomes and therefore require fewer, but more carefully planned and protected, occasions of assessment. Thus, assessment is best resourced and deployed to assure significant outcomes, with grade-bearing tasks loaded towards the end of a program, aligned with explicit learning outcomes, particularly at the program level.

This means avoiding graded assessment tasks that require responses to common questions or problems which can be addressed through simple searching or derived via genAI. Such tasks can however be useful formatively for student self-checking prior to more complex and challenging assurance tasks that are highly contextualised, interactive, synchronous, personalised or part of a unique sequence.

Design feedback sequences to support learning

Assessment for learning should be used in early parts of a unit or course of study, and in early years of a program. For students to learn from undertaking such tasks, they should also receive improvement-focused comments. They then have opportunities to act on this information and demonstrate improved understanding in subsequent work. Such tasks could authentically incorporate genAI. Specific, forward-looking, contextualised feedback comments from teachers, peers or others (including genAI) can prompt student learning as well as promote effective and appropriate engagement with genAI.

Develop student capability to identify 'what good looks like' through assessment

With the rise of genAI, students will need to make more judgements about what 'good looks like' with respect to others' work, especially AI-produced artefacts. Assessment practices should therefore develop students' evaluative judgement, that is, their ability to judge the quality of their work and others'. Assessment approaches for this purpose may include students identifying markers of quality in exemplars; critiquing AI products; co-developing assessment criteria; and making comparisons of completed work and engaging with standards/rubrics through peer and self-feedback. Assessment tasks for this purpose can require students to demonstrate or articulate their own evaluative judgements.

Devise multiple submission formats to make assessment more equitable

Assessments should *only* discriminate on the relevant criteria, rather than judge qualities of students irrelevant to the standards required to be met, such as the task format. Including a variety of submission formats for a task, and/or across a course or program may therefore improve equitability of assessment. This can balance out whether the use of genAI itself becomes a barrier (for example due to cost) or an enabler (for example, mitigating the impact of personal circumstances).

Focus on evidencing that outcomes are met

Rather than adopting a cheating or academic integrity framing of genAI, we suggest an assessment framing, focused on evidencing outcomes. For most outcomes this will require a range of different types of tasks in different circumstances that can provide different types of evidence. For example, an outcome related to problem solving might be evidenced through a mixture of interactive oral assessment of groupwork, supervised online exams, practicum reports, and logs from ChatGPT sessions. Rather than trying to develop a single task that is impervious to genAI, we instead suggest a focus on sequences of tasks (within or across units) that collectively provide greater assurance that outcomes have been met. Doing this will require a clear understanding of the outcomes being assessed and it may not always be possible to disentangle one outcome from another especially in complex tasks.

Adapting current assessment practices to account for genAI

Have open conversations about genAI with students

We need to consider what stance we take in relation to genAI in assessment. If we assume that students are going to use it to cheat, then we risk further fracturing student-staff relationships. However, if we have open conversations about how students might use genAI (in sanctioned ways) and what may or may not breach academic integrity rules (and why) we can build trust. We should acknowledge the use of genAI in our own teaching, assessment and research, thus modelling appropriate and effective use. We want students to build productive, safe, ethical and critical relationships with this and future technologies. By having such conversations with students, we might help tackle uncertainties about what is acceptable, how to acknowledge the sources drawn on, and the limits of genAI.

Review rubrics and other forms of assessment criteria

Rubrics and assessment criteria communicate expectations regarding the quality of student work. Criteria for success that rely on presentation of common views or knowledge recall, are (further) called into question by genAI. Such criteria should be reduced in weight or even eliminated. Sometimes they can be integrated in other criteria that match to demonstrations of critical thinking, originality and so on. Where such criteria reflect lower-level learning outcomes that must be met, this will require high-stakes, summative assessment approaches that restrict student use of genAI, such as interactive oral assessment or synchronous in-class work.

Specify assessment situations where it is appropriate or inappropriate to use genAI

In many disciplines there are learning outcomes that students need to be able to demonstrate without the use of genAI. However, once students have demonstrated these outcomes to the required standard, is there a need for them to continue to do so for the rest of their degree? The answer will be different for different outcomes, but for some, allowing the use of genAI will enable students to 'offload' time consuming aspects of a task that are less educationally relevant, and spend more time on other aspects that matter most.

Design tasks to promote students' portrayal of their unique achievements

In many assessment tasks students produce the same work as each other– but this is a practice that may usefully be challenged. Originality of thought and work will be an increasing focus in a world of AI as genAI relies on predictive patterns in data. Therefore, so long as tasks address the necessary learning outcomes, they should permit differences in student response, rather than eliciting what they have in common. For example, use tasks that involve students positioning their work in their own life, work or community with opportunities for choice in portrayal. There are caveats: originality might not be expected in introductory classes and students must be scaffolded to enable them to both produce and articulate what is different.

Develop and assess critical digital literacies

GenAI highlights the need to assess critical digital literacies. In the first instance, students need to learn about appropriate use of AI aids. Students also should learn about the ethical complexities associated with these new technologies. For example, genAI compositions may privilege certain repositories and ways of knowing or promote discriminatory practices as they reflect the sources that they draw on. Likewise, students need to know that sharing information about themselves may contribute to corporate interests in ways they do not want. Courses need to model critical digital literacies, but we also cannot assume learning through observation and use: we should flag the seriousness of 'need to know' critical digital literacies through learning outcomes and associated assessment tasks.

See CRADLE suggests series for practical summaries of this work.

Cite as: Bearman, M., Ajjawi, R., Boud, D., Tai, J. & Dawson, P. (2023). CRADLE Suggests... assessment and genAI. Centre for Research in Assessment and Digital Learning, Deakin University, Melbourne, Australia. doi:10.6084/m9.figshare.22494178