

+ Assessment and feedback case study compendium 2026

Creating authentic assessments for an AI-enabled world

Edited by Vic Stephenson, Senior Consultant & Stuart Norton,
Head of Educational Excellence, Advance HE



Contents

Advancing authentic assessment: development of an institutional framework and guide for authentic assessment	4
Dr Michelle Cordingley, Professor Jackie Potter and Dr Katharine Welsh, University of Chester, UK	
Ctrl + Alt + Crit: re-imagining assessment and feedback with Generative AI in Art, Design and Creative disciplines	11
Louise O'Boyle, Ulster University, UK	
The challenges of assessment incorporating both AI literacy and criticality: lessons from an AI module at the University of South Wales	17
Liam Newton, University of South Wales, UK	
Embedding work-integrated learning through human-centred design for an AI-driven environment	22
Nina Edward, Wai Hoe Yan and Yizhi Wu, University of Sydney, Australia	
Flawed by design: embedding AI to drive authentic assessment and integrity	30
Rania Shibl, Southern Cross University, Australia, and Dr Judy Watson, University of the Sunshine Coast, Australia	
When the threat becomes a tool: using generative AI to create authentic assessments for final year Life Sciences students	35
Dr Emma Reid, Lecturer, University of Glasgow, UK	
Transforming the university learning experience and enhancing student success through authentic assessment and feedback	40
Frantzeska Kolyda, University of Westminster; Johnathan Lucky, Sky; Helena Disu, University of Westminster, UK	
Making research reasoning visible: redesigning a final-year literature review for a GenAI-enabled world	46
Denise Hough, Victoria Paterson and Barbara K Mable, University of Glasgow, UK	

Advancing authentic assessment: development of an institutional framework and guide for authentic assessment

Dr Michelle Cordingley, University Innovation Fellow; Professor Jackie Potter, Dean of Academic Innovation and Development and Dr Katharine Welsh, University Innovation Fellow, University of Chester, UK

Background

Improving assessment practices is essential for enhancing learning and teaching within higher education (Advance HE, 2024). A Better Assessment Strategic Innovation Forum (SIF) approach brought together staff from across the University of Chester to critically explore existing assessment practices and identify areas for improvement that could enhance student continuation, completion and progression, while also supporting positive student and staff wellbeing (Potter et al, 2024). Through structured engagement and collaborative inquiry conducted between January and June 2023, this wide-reaching community-led process resulted in the generation of a 10-point institutional action plan for delivery in 2023-24 academic year. As highlighted by Advance HE (2024), higher education institutions play a critical role in advancing, modelling and disseminating innovative assessment practices. One of the key actions emerging from the Better Assessment SIF process was therefore to develop a Guide to Good Assessment.

Although authentic assessment has been widely used in higher education for more than 30 years, more recent scholarship has focused on clarifying and defining the concept (Ajjawi et al, 2023; Kandlbinder, 2025; McArthur, 2023). As authentic assessment is a fundamental principle of the University of Chester Education Plan, the guide was subsequently renamed A Guide to Authentic Assessment (Cordingley, 2024a). The guide was designed to foster continuous learning and innovation among staff and to promote effective implementation of authentic assessment within teaching practice. To provide practical illustrations of authentic assessment in action, an accompanying document, Authentic Assessment Examples of Practice (Cordingley, 2024b), was compiled with contributions from colleagues across the institution to complement the guide.

Unbeknown at the time of the Better Assessment SIF, shortly after its conclusion, the University of Chester introduced a new academic framework and institution-wide curriculum, the Future Skills Curriculum, in which continuous, embedded authentic assessment is a key feature. This required course teams across the institution to engage in the design and redesign of courses and associated assessment strategies for the 2025-26 academic year. The Guide to Authentic Assessment therefore became a timely and practical resource to support course teams during the curriculum design sprint process. In doing so, the guide has helped ensure alignment between strategic institutional change, curriculum

development and day-to-day assessment practice, supporting positive and sustainable change in assessment approaches. This case study shares insights from the process of developing the University of Chester Guide to Authentic Assessment and an underpinning institutional framework for authentic assessment.

Approach

The University of Chester Guide to Authentic Assessment (Cordingley, 2024a) was developed over a nine-month period by members of the University of Chester Centre for Academic Innovation and Development (CAID). The work was led by Dr Michelle Cordingley (University Innovation Fellow), with support from Professor Jackie Potter (Dean of Academic Innovation) and Dr Katharine Welsh (University Innovation Fellow), and the guide launched in May 2024.

The development process included creating an institutional authentic assessment framework, producing the guide, identifying and curating examples of practice from across the University, and disseminating the guide to support staff engagement and facilitate implementation.

Authentic Assessment Framework

With authentic assessment as a fundamental principle of the University of Chester Education Plan, the first step in the process was to develop an institutional framework (Figure 1). Existing models, particularly the QAA Inclusive Assessment Attributes (QAA, 2022) and the Jisc Principles of Good Assessment and Feedback (Knight and Ferrell, 2022), were reviewed and Self Determination Theory (Ryan and Deci, 2000) was also considered. From this review, 13 key sub-themes relevant to assessment practices at the University of Chester were identified.

These sub-themes were then organised into three overarching themes, providing a coherent framework to support the design of authentic assessments that:

- + foster motivation to learn by building autonomy, competence and relatedness
- + supports students' futures
- + promote positive wellbeing through inclusive and compassionate practices.



Figure 1. University of Chester Authentic Assessment Framework

Guide to Authentic Assessment

The development of the Guide to Authentic Assessment (Cordingley, 2024a) began by articulating the importance of authentic assessment for student learning and introducing the University of Chester Authentic Assessment Framework. The guide then explores each of the three main themes in more detail, including key resources. The remainder of the guide is structured around the 13 supporting sub-themes framed using the phrase '*Authentic assessment works best when...*' Generative artificial intelligence (GenAI) was employed to help identify key considerations for each sub-theme, which were then reviewed and adapted to ensure relevance within the Chester context. Each subtheme includes a dedicated section outlining essential considerations, along with links to relevant examples of practice that illustrate how the sub-theme can be applied in practice.



Figure 2. Illustrative pages from the Guide to Authentic Assessment (Cordingley, 2024a)

Authentic assessment example of practice

To provide practical illustrations of authentic assessment in action, an institution-wide call for examples of practice was launched. Colleagues were invited to submit contributions using a templated form, providing details on context, impact, practical tips and key resources. Contributors were also asked to indicate which sub-themes their examples exemplified. One of the initial challenges was gathering examples of practice, as time-constrained staff and a first call-out yielded few responses. A second, more targeted, call encouraging colleagues to highlight areas of good practice within their Schools and Divisions, and to identify people to approach, proved more successful. A peer review process was implemented to enhance the accessibility and usability of the examples of practice with contributors participating in reviewing others' submissions and reporting that this process helped improve their own contribution. In total, 18 examples of practice were gathered from across the University's three faculties. These examples were compiled into a single document – Authentic Assessment Examples of Practice (Cordingley, 2024b) – organised under the three main authentic assessment themes, with the primary theme clearly highlighted and relevant sub-themes indicated (Figure 3).

Assessment and feedback case study compendium 2026

Creating authentic assessments for an AI-enabled world

Example of Practice 4

Commercial Law students 'thinking commercially' for enhanced employability

Andrea Todd
andrea.todd@chester.ac.uk
LAS007 Commercial Law
LLB Law
School of Law and Social Justice
Faculty of Arts, Humanities and Social Sciences

At a Glance
Law students are reminded at every turn of the necessity to think commercially if they wish to land – and succeed in – a graduate legal role. However, at undergraduate level little attention is generally afforded to ensuring students understand what commercial awareness, and thinking commercially, mean.

This case study looks at how engaging with Commercial Law via a client case study simulation – and being assessed by way of practical tasks commonly encountered by Commercial solicitors – has enabled students to:

- grasp the fundamental concepts of Commercial Law;
- appreciate the importance of understanding, and tailoring their legal advice to, a client's commercial priorities, and
- practice the skills required in the world of work.

This module's approach to teaching, learning and assessment provides students with the opportunity to frame their academic learning within the context of the mindset, and skillset, that will be expected of them in the world of work.

Practical Tips for Assessment Success

- Allow plenty of time in class for practising the practical skills for assessment.
- Foster a supportive environment during workshops in which peer assessment and feedback can be constructively given and positively received.
- Use real world documents (e.g. attendance notes, emails from clients, contracts) to replicate documentation encountered in the world of work.

Assessment Supports Students' Futures

Assessment demonstrates student's skills and competencies

Assessment enables dialogue around learning

Assessment delivers high quality feedback

Assessment encourages collaborative learning

Assessment promotes positive student wellbeing

In More Detail

Context

Knowledge of the law is taken as read in the commercial world. Any interview with a graduate recruiter, and any Legal 500 entry, will impress upon the reader the most important feature of good legal advice: having a lawyer who is commercial in their approach.

A focus on 'commerciality' is lacking from undergraduate legal education. The Commercial Law module has been designed to fill this gap, providing students with the opportunity to 'think commercially' by providing advice tailored to the client's specific needs and priorities and to practice the skills needed in the workplace to enable them to succeed in a commercial law environment.

Via a series of eight workshops, the module follows a start-up business from its inception, with students advising the client (the 'Workshop Client') as the sessions progress on the choice of business medium to use to set up the business, the choice of funding options to expand their operations, dealing with the challenges of buying and selling as a business, and dealing with commercial agents to extend the geographical scope of business operations. The module leader has drawn on her experience of working as a solicitor in a commercial practice to pull together the Workshop Client's journey as a young business and the challenges it encounters.

Assessment Design

Assessment in this module is by way of two practical tasks:

1. A recorded oral presentation, with accompanying slide deck, to two previously unknown clients (separate from the Workshop Client) regarding choice of business medium and the types of shares to allot to an incoming investor, and
2. A due diligence report for a further new client, revealing the legal problems unearthed during investigations into a client's proposed new business acquisition and addressing, in practical terms, how the client can protect themselves from the legal and financial risk associated with these problems.

A benefit of assessment tasks tailored to the specific concerns, priorities and attitudes of a particular client is that a sound performance requires not only strong legal knowledge but also the exercise of individual judgment, which is difficult to produce by way of artificial intelligence.

Workshop Tasks

At the start of every workshop, students are provided with a new set of instructions from the Workshop Client which take them one step further in their journey as a young business. Students spend time during the workshop working in groups to devise legal solutions to the challenges faced by the client, and then put together an exposition of their advice by way of oral presentation or written report, replicating what is required for the assessment tasks.

Students receive peer and module leader feedback on their legal answers as well as their skills practised during every workshop. In doing so, the

"The assessment helped me to focus on how to address clients, the correct terminology to use, and how to display and deliver a presentation in a professional way – all skills that will utilise in the future"

LAS006 Student Feedback

module leader invites each group of students to present their work from the front of the room and sit down with the remaining students to watch the presentation. The module leader then invites constructive feedback from the floor (having explained the criteria we are looking for when making the assessment) and contributes to discussions not as the leader of the session, but as a member of the floor. It is worth noting that sessions have been elongated over recent years to deal with feedback that some students had sometimes felt rushed in the preparation of their work for exhibiting to peers, and/or had felt that they had not had sufficient opportunity to deliver their work and/or to give feedback. As such, each session is now two hours long with at least a third of the session dedicated to presenting, and receiving feedback on, the legal content and skills practised.

Impact

Students report benefiting from having a context in which to place their legal knowledge, thus improving engagement and encouraging successful academic outcomes:

- 'My grades are higher and my understanding is stronger'
- 'This module has emphasised that knowing the law is not enough, but being able to apply it to practical everyday situations has really helped me build my skills for my future career as a solicitor'
- 'The module has given us a glimpse into situations we might be faced with in the future, and provided us with opportunities to apply what we have learned in a practical way'

Students also report the benefits of the module from an employability perspective:

- 'The Commercial Law module was so helpful to me when I did a work placement this summer. Having an awareness of how to word my advice (like we did in our presentation and in the due diligence report) was so useful when I was communicating research to partners at the firm. I think the module really helped me with professionalism and the ability to communicate clearly'
- 'Having the confidence to advise a client (once I start work) has been greatly helped by the presentation task'
- 'The assessment helped me to focus on how to address clients, the correct terminology to use, and how to display and deliver a presentation in a professional way – all skills that I will utilise in the future'

Insights

Lawyers in practice are not required to write essays. The approach to teaching, learning and assessment used in this module enables students (who may not excel in traditional assessments) to shine by undertaking the tasks that commercial lawyers do in practice. This leads to both an increased sense of confidence and to enhanced employability prospects.

Using real world documentation and providing assessment by way of commonly undertaken tasks from the world of work, and encouraging students to provide peer feedback on skills during class time, is replicable to other modules, regardless of discipline.

Further Resources

M. van Zundert, D. Sluitermans, J. van Marrewijk (2010), *Effective peer assessment processes: Research findings and future directions*, *Learning and Instruction*, 20(4), 270-279. <https://doi.org/10.1016/j.learninstruc.2009.05.005>

V. Murphy, J. Fox, S. Freeman, & N. Hughes (2017), 'Stepping It Real? A Review of the Benefits, Challenges and Steps Towards Implementing Authentic Assessment: The All Ireland Journal of Teaching and Learning in Higher Education (ATHE-IJ) 9. <https://www.athe.ie/assets/pdfs/2017-athe-journal-vol9.pdf>

Figure 3. Illustrative authentic assessment example of practice. Taken from Cordingley, 2024b

Outcomes

The key outputs of this work were the Guide to Authentic Assessment (Cordingley, 2024a) and Authentic Assessment Examples of Practice (Cordingley, 2024b), launched institution-wide in via TLInnovate, the University of Chester's online teaching and learning community of practice. The launch event introduced the guide and featured several contributors presenting their examples of practice, with more than 120 participants from across the institution in attendance.

Following the launch, more than 100 course teams actively applied the Authentic Assessment Framework and used the resources to reimagine assessment during a curriculum design sprint process. A radar tool, based on the framework, was also developed as a visual reflective aid, enabling course teams to evaluate their assessment design across the multiple sub-themes and facilitate discussion on potential improvements.

Additionally, an interactive workshop 'Don't reinvent the wheel – spin the authentic assessment wheel' was developed for the University of Chester Teaching and Learning Conference 2025. The workshop, attended by academic and professional services staff from across the institution, supported reflective and collaborative development of authentic assessments using the framework and a set of prompt questions related to the sub-themes.

The initiative demonstrated clear impact in three areas:

+ Improved student learning opportunities

“I will look to include more opportunities for formative assessment and peer dialogue around the assignment”

+ Increased staff confidence

“Design an assessment brief for a new type of assessment”

+ Reflexive practice

“Consider my assessments and tweak further using the prompt questions”

The framework and guide have also been disseminated externally, including presenting a poster at the Advance HE Assessment and Feedback Symposium 2025 and via a mini keynote focused on Self-Determination Theory and authentic assessment at the Assessment in Higher Education AHE Conference 2025. An interactive workshop was also delivered at the Compassionate Assessment Practice Sharing Symposium where attendees reflected on their own practice using prompt questions based on the ‘authentic assessment promotes positive wellbeing’ theme of the University of Chester Authentic Assessment Framework. Feedback from the external activities illustrates impact on practice beyond Chester:

“The guide is very inspiring. I can see how the language and ethos here is transferable to lots of different teaching contexts. It has inspired me to develop a framework for a new course which I am involved in.”

“I have found the whole document on authentic assessment a real source of help and inspiration as we design a new programme.”

The University of Chester Authentic Assessment Framework and the Guide to Authentic Assessment can be used directly to inform and enhance authentic assessment in individual practice at Chester and beyond, or the process outlined in this case study can be adapted by others to develop an institution-wide framework with aligned guides and supporting resources.

References

- Advance HE (2024) *Essential Frameworks for Enhancing Student Success: Enhancing Assessment in Higher Education*. York: Advance HE. Available at: advance-he.ac.uk/guidance/teaching-and-learning/essential-frameworks-enhancing-student-success [accessed 1 February 2026].
- Ajjawi, R, Tai, J, Dollinger, M, Dawson, P, Boud, D and Bearman, M (2023) 'From authentic assessment to authenticity in assessment: broadening perspectives', *Assessment and Evaluation in Higher Education*, 49 (4): 499-510.
- Cordingley, M (2024a) 'A Guide to Authentic Assessment'. Chester: University of Chester. Available at: www.chester.ac.uk/centre-for-academic-innovation-and-development/ [accessed 1 February 2026]
- Cordingley, M (2024b) 'Authentic Assessment Examples of Practice'. Chester: University of Chester. Available at: www.chester.ac.uk/centre-for-academic-innovation-and-development/ [accessed 1 February 2026].
- Kandlbinder, P (2025) 'Signature concepts of authentic assessment', *Higher Education Research and Development*, 1-15.
- Knight, S and Ferrell G (2022) *Principles of good assessment and feedback*. Bristol: Jisc. Available at: www.jisc.ac.uk/guides/principles-of-good-assessment-and-feedback [accessed 1 February 2026].
- McArthur, J (2023) 'Rethinking authentic assessment: work, well-being, and society', *Higher Education*, 85 (1): 85-101.
- Potter, J, Falcon, R, Collins, R and Simpson (2024) 'Delivering 'better assessment' through a whole institution, community-led strategic innovation approach', in Norton, S and Stephenson, V (eds) *Enhancing assessment and feedback: a case study compendium*. York: Advance HE, pp 61-65. Available at: s3.eu-west-2.amazonaws.com/assets.creode.advancehe-document-manager/documents/advance-he/Enhancing_assessment_and_feedback_case_study_compendium_1729083820.pdf [accessed 7 February 2026].
- Ryan, R M and Deci, E L (2000) 'Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being', *American Psychologist*, 55 (1): 68-78.
- QAA (2022) 'Inclusive assessment Attributes'. Gloucester: QAA. Available at: www.qaa.ac.uk/docs/qaa/members/inclusive-assessment-attributes.pdf [accessed 1 February 2026].
-

Ctrl + Alt + Crit: re-imagining assessment and feedback with Generative AI in Art, Design and Creative disciplines

Louise O'Boyle, Associate Dean, Academic Quality and Student Experience, Faculty of Arts, Humanities and Social Sciences, Ulster University, UK

Background

By late 2024, generative AI (GenAI) had become embedded in students' everyday assessment practices. HEPI's Student Generative AI Survey 2025 revealed that 92% of UK undergraduates had used AI tools, with 88% reporting their use to support assessed work. Yet only a third felt their institution supported them to use these tools effectively and ethically (Freeman, 2025).

Simultaneously, sector bodies were publishing guidance beyond misconduct warnings. The Quality Assurance Agency for Higher Education (QAA) emphasised reviewing assessment structurally (2024), and the Russell Group framed AI literacy as a graduate expectation (2025). This shift was reflected in Australia's Tertiary Education Quality and Standards Agency (TEQSA) advice to reduce reliance on easily automated formats (Tertiary Education Quality and Standards Agency, 2023).

Across Art, Design and Creative disciplines (ADC), these signals mirrored staff concerns. This case study emerges from situated knowledge. As Associate Dean, Chair of GLAD and Expert Trustee of CHEAD, I experience GenAI's impact operationally in integrity panels, curriculum development and cross-institutional conversations. Through GLAD events (2023; 2024; 2025), I have witnessed automation, standardisation and output-focused assessment manifest as sector-wide concerns. These roles afford privileged learning but also responsibility: to translate cross-institutional insight into frameworks supporting colleagues navigating these challenges. I speak from ADC education in a post-1992 context, learning alongside those I support.

From these discourses, three persistent tensions emerged.

- + First, automation versus authenticity. When students can generate competent artefacts at speed, traditional inferences from product to capability become less secure.
 - + Second, standardisation versus speculation. GenAI excels at producing 'competent normal' by averaging aesthetics. Yet ADC disciplines reward divergence and speculative futures. If rubrics reward polished normality, GenAI will meet the criteria.
 - + Third, output versus process. Institutions encourage students to show their process, but marks still cluster around finished artefacts. In a GenAI context, process is often the most credible site of learning, yet it remains under-assessed.
-

Alongside these tensions sit pragmatic concerns: cognitive disengagement when students outsource ideation; hallucination; and encoded bias raising questions of representational politics. Assessment must therefore shift from ‘is this a good poster or essay?’ to ‘can this student demonstrate awareness, judgement, authorship and ethical stance?’

None of these tensions were new, GenAI simply intensified them. Neither a ‘ban it and hope’ nor an uncritical ‘it’s just another tool’ stance would suffice. What was missing was a framework operationalising critical GenAI literacy into assessable criteria. The **Ctrl + Alt + Crit** framework emerged in response to that need.

Approach

The title plays on the ‘Ctrl+Alt+Del’ keystroke reset, but instead of ‘Delete’, I propose ‘Critique’. The provocation is simple, instead of shutting the system down in panic, we reassert control through structural assessment design (Ctrl), explore alternatives to binary positions (Alt), and treat critique as a core academic function (Crit).

The framework was shaped through three strands: a focused literature and policy review; qualitative feedback gathered through routine quality enhancement conversations; and cross-programme analysis from my role positionality, rather than a formal research study.

The literature strand drew on evaluative judgement and students’ capability to appraise quality (Tai et al, 2018; Orr and Shreeve, 2019; Mulholland, 2019), including AI-generated material (Hoover, 2024; Rodgers, 2025), and feedback literacy (Carless and Boud, 2018). It also engaged critiques of the ‘enforcement illusion’, where tasks are labelled AI-permitted or AI-prohibited while their underlying design remains unchanged (Spruijt, 2024).

The staff strand surfaced practical pressures: worries about integrity, workload and fairness, alongside recognition that graduates will need to work with GenAI professionally (Microsoft and LinkedIn, 2024). Staff described being ‘pushed’ into critical AI guide roles as they field queries from students as to which tools are permissible to use, how do they avoid reproducing bias and where is the line between enhancement and outsourcing – pedagogic and disciplinary questions as much as technical ones.

Validation through GLAD events suggested these tensions were systemic to ADC rather than localised to any single institution (GLAD, 2023; 2024; 2025). ADC disciplines can act as a test site for the wider sector, training students to read conditions, disrupt norms and propose credible alternatives (Abegglen et al, 2024; Rodgers, 2019). Enabling critical pedagogy requires attention to power, voice and transformative learning (Seal and Smith, 2021). GenAI excels at remixing what exists; our students need to generate what does not yet exist.

Bringing these strands together, I sketched an initial version of **Ctrl + Alt + Crit** and refined it through critique from trusted colleagues. The framework distils into three design imperatives:

Ctrl (Structural Control) reasserts control through assessment design at programme level, shifting attention from surveillance of students' tool use to the structure of tasks, sequences and criteria.

Alt (Alternative Configurations) moves beyond 'AI is banned' versus 'AI is allowed' towards specific, named configurations of human-AI interaction, such as GenAI as provocateur, drafting partner, simulated client or feedback source to be evaluated.

Crit (Centring Critique) makes critique of GenAI outputs, systems and institutional arrangements examinable.

From this work, five pragmatic assessment strategies were drawn and illustrated below.

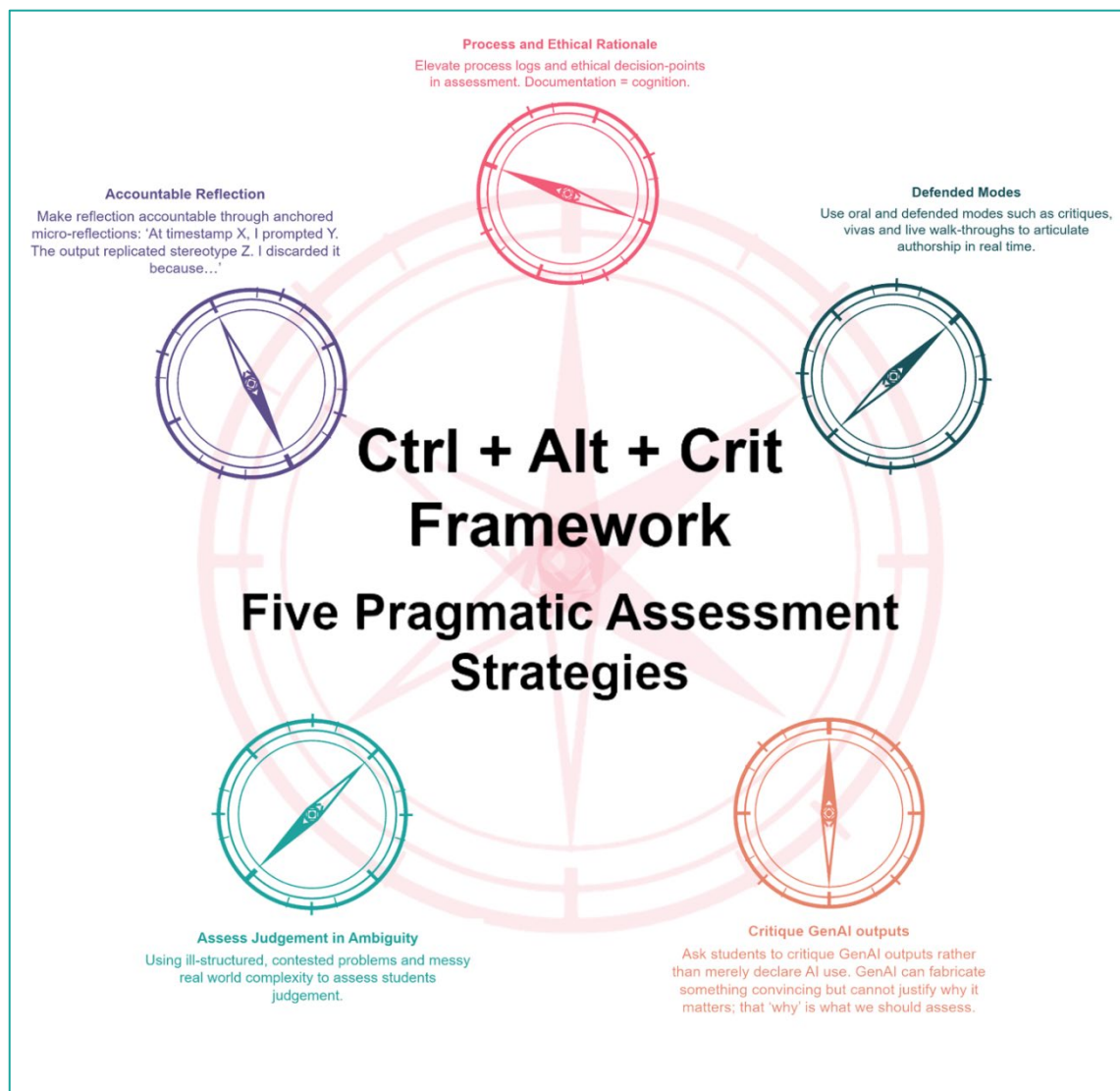


Figure 1. Ctrl + Alt + Crit Framework Illustration (Source: O'Boyle, 2025)

Outcomes

This case study is provocation, not solution. **Ctrl + Alt + Crit** is offered as a developmental framework refined through colleague critique and sector validation, but not yet formally piloted. The next phase involves structured implementation with ADC programme teams, evaluating how the framework shapes assessment redesign and whether the design imperatives prove tractable. The honest position is uncomfortable: GenAI has exposed longstanding inadequacies in how and what we assess.

The three tensions identified earlier are not new problems caused by GenAI; they are problems GenAI has made impossible to ignore. Assessment has drifted towards standardisation and easily marked outputs. We have designed tasks that a sophisticated pattern-matching system could approximate. GenAI is that system; its arrival is a mirror.

This demands we reframe assessment and feedback. Instead of gatekeeping, assessment becomes a compass orienting students: where they are, where the discipline is, where they could go next. Feedback becomes provocateur, disrupting in ways that sharpen intent.

The sector faces a choice. One path leads to elaborate detection, surveillance and policy-patching, treating GenAI as threat while leaving underlying structures intact. This path will fail as detection tools lag behind generative capabilities, and policy clauses may proliferate without impact on practice.

The other path is harder but more honest. We must revisit not just tasks but curricula, learning outcomes and pedagogies. What capabilities do we actually want graduates to develop? Which are distinctively human, and which have we assessed simply because they were convenient proxies for learning?

Yet this discomfort is precisely where intellectual excitement should begin. ADC disciplines, with their traditions of critique and making-as-thinking, are well placed to model what assessment might become: developing human judgement in dialogue with powerful tools (Hern, 2025; Hoffman and Beato, 2025). When we design assessment in ADC, we assess whether students can read and shape culture under algorithmic conditions. That is employability (World Economic Forum, 2023), but also democratic capacity in an AI-mediated public sphere.

The sector cannot policy-clause its way out of this moment. GenAI offers an opportunity to redesign assessment so it cultivates human capability rather than rewarding outputs machines can produce. That redesign requires resources, protected time, and tolerance for productive failure. **Ctrl + Alt + Crit** is offered in that spirit. Not as an answer, but as an invitation to stop burying our heads in the sand and start exploring new territory together.

Open Access Statement

For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any author-accepted manuscript version arising from this submission.

References

- Abegglen, S, Nerantzi, C, Martínez-Arboleda, A, Karatsiori, M, Atenas, J, Rowell, C (2024) *Towards AI Literacy: 101+ creative and critical practices, perspectives and purposes*. Calgary: #creativeHE. Available at: doi.org/10.5281/zenodo.11613520
- Carless, D and Boud, D (2018) 'The development of student feedback literacy: enabling uptake of feedback', *Assessment and Evaluation in Higher Education*, 43 (8): 1315-1325.
- Freeman, J (2025) *Student Generative AI Survey 2025*. Oxford: HEPI. Available at: www.hepi.ac.uk/reports/student-generative-ai-survey-2025/ [accessed 1 March 2025].
- GLAD (2023) *AI and the art school*. Available at: www.gladhe.com/home/aiartschool [accessed 15 January 2025].
- GLAD (2024) *GLAD Conference 2024*. Available at: www.gladhe.com/home/2024conference [accessed 15 January 2025].
- GLAD (2025) *GLAD Symposium 2025*. Available at: www.gladhe.com/home/2025symposium [accessed 1 September 2025].
- Hern, A (2025) 'Reid Hoffman: start using AI deeply – it is a huge intelligence amplifier', *The Guardian*, 22 March. Available at: www.theguardian.com/technology/2025/mar/22/reid-hoffman-superagency-start-using-ai-deeply-it-is-a-huge-intelligence-amplifier [accessed 23 March 2025].
- Hoffman, R and Beato, G (2025) *Superagency: what could possibly go right with our AI future*. New York: Simon and Schuster.
- Hoover, A (2024) 'Students are likely writing millions of papers with AI', *WIRED*, 9 April. Available at: www.wired.com/story/student-papers-generative-ai-turnitin/ [accessed 1 July 2025].
- Microsoft and LinkedIn (2024) *AI at work is here. Now comes the hard part. Work Trend Index Annual Report*. Microsoft and LinkedIn. Available at: www.microsoft.com/en-us/worklab/work-trend-index/ai-at-work-is-here-now-comes-the-hard-part [accessed 1 June 2024].
- Mulholland, N (2019) *Re-imagining the art school: pedagogy and artistic learning*. 1st edn. Switzerland: Palgrave Pivot.
- Orr, S and Shreeve, A (2019) *Art and design pedagogy in higher education: knowledge, values and ambiguity in the creative curriculum*. 1st edn. Abingdon: Routledge.
- QAA (2024) *Quality compass: navigating the complexities of the artificial intelligence era in higher education*. Gloucester: Quality Assurance Agency. www.qaa.ac.uk/docs/qaa/news/quality-compass-navigating-the-complexities-of-the-artificial-intelligence-era-in-higher-education.pdf [accessed 1 March 2025].
-

Rodgers, P (2019) Design School: after boundaries and disciplines. USA: Vernon Press.

Rodgers, R (2025) 'I'm not convinced ethical generative AI currently exists', WIRED, 20 February. Available at: www.wired.com/story/the-prompt-ethical-generative-ai-does-not-exist [accessed 1 March 2025].

Russell Group (2025) Russell Group principles on the use of generative AI tools in education. London: Russell Group. Available at: www.russellgroup.ac.uk/policy/policy-briefings/principles-use-generative-ai-tools-education [accessed 1 March 2025].

Seal, M and Smith, A (2021) Enabling critical pedagogy in higher education. 1st edn. St Albans: Critical Publishing.

Spruijt, J (2024) 'Ethical considerations regarding the use of AI in higher education', Open Innovation, 24 June. Available at: openinnovation.eu/24-06-2024/ethical-considerations-regarding-the-use-of-ai-in-higher-education/ [accessed 4 July 2025].

Tai, J, Ajjawi, R, Boud, D, Dawson, P and Panadero, E (2018) 'Developing evaluative judgement: enabling students to make decisions about the quality of work', Higher Education, 76: 467-481. Available at: doi.org/10.1007/s10734-017-0220-3

Tertiary Education Quality and Standards Agency (2023) Assessment reform for the age of artificial intelligence. Melbourne: Tertiary Education Quality and Standards Agency.

World Economic Forum (2023) Future of Jobs Report 2023. Geneva: World Economic Forum. Available at: www3.weforum.org/docs/WEF_Future_of_Jobs_2023.pdf [accessed 1 March 2024].

The challenges of assessment incorporating both AI literacy and criticality: lessons from an AI module at the University of South Wales

Liam Newton, Senior Lecturer in Business and Management, University of South Wales, UK

Background

Assessments can no longer ignore the existence of generative AI tools, or their use by students. This case study relates to the development of a module focusing on AI literacy for BA (Hons) Business and Management students at the University of South Wales, in particular focusing on the challenge of developing authentic assessments for the module that test both students' ability to use generative AI, as well as their criticality and reflexivity in the use of those tools.

This credit-bearing module was first designed and added to the course for the 2024-25 academic year, in response to two trends noted by staff on the degree: firstly, that graduate employers within the field of business and management are increasingly demanding AI literate graduates, and secondly, that although students were commonly using AI technologies, they were often using tools poorly, or unknowingly using the wrong tools for the job.

Graduate employability is one of the central aims of this Business and Management degree – and for many UK HEIs. This is due to the growing importance of metrics such as the Graduate Outcomes Survey, and to a range of stakeholders, including students and the UK government, seeing employability as a key function of higher education (Cheng et al, 2022). For Business and Management graduates, AI literacy is increasingly becoming a differentiating factor in employability (Portocarrero Ramos et al, 2025) – meaning that it is no longer an option to ignore generative AI technologies, but that responsible management education focused on vocational outcomes for students must instead grapple with how to develop AI literate graduates (Kennedy, 2024).

For this course, the importance of developing AI literacy was underlined by anecdotal observations from the course team around students' existing use of AI. Many students now use AI to assist them with their studies or assessments – a trend observed in many institutions in the UK (Arowosegbe et al, 2024) and internationally, but when questioned as to *how* they had used AI, the answers betrayed a lack of AI literacy.

Students on this course had used AI tools not designed for the tasks they had attempted, not understood the limitations of the tools they had used, and often admitted that their habits around AI use had come through talking to other students, rather than developing their own understanding of the tools themselves.

This situation did not benefit either the students or staff: academics were frustrated at students using AI in a way they considered to be misconduct, while many students were missing out on developing competence in both the course content and the use of AI tools. This module, therefore, aimed to address this through the explicit development of AI literacy – including both technical AI literacy and cognitive literacies around reflexive, thoughtful and ethical use of AI.

With the overall module purpose and fit in mind, the challenge then became how to effectively assess this module – designing assessments that tested both students' digital and cognitive literacies around AI in a rigorous and authentic manner, and ensuring that the assessments would remain valid despite the rapid advancement of AI technologies.

Approach

A new assessment format was decided for this module. The previous module this replaced had a more traditional assessment mix, with students required to do a 3,000 word written report. This wasn't thought to be the best approach for this new module: it would be difficult to test students' practical AI literacy through only a written report, and growing levels of AI mis-use meant this format posed an increasing challenge when trying to assess students' criticality and reflective ability. It was also important for the new assessment format to fit within existing workload constraints, particularly in order to be scalable without significant extra institutional investment. As a result, two assessments were decided for this new module: a portfolio of AI-created outputs with reflections and process logs, before taking part in a novel 'Professional Discussion' assessment designed to test their awareness and critical thinking ability.

The first assessment, the portfolio, required students to use AI tools to generate three distinct outputs suitable for use by businesses, along with a reflection on the suitability of each output and a log of the process undertaken. It was required that each output be created using a different generative AI tool, and be fundamentally different from each other (for example, students were only allowed to produce one long-form text response). This assessment was designed to be authentic, with outputs expected to be immediately suitable for use in a business context, while also promoting criticality through the reflections.

The requirement for three different outputs also had the intended learning outcome of building students' AI literacy – the assessment rubric rewarded innovation and novel uses of AI, reflecting the importance of entrepreneurial and innovative behaviours in business and management studies (QAA, nd). While almost all students entered the module capable of using AI tools to generate long-form text, very few had considered the range of possibilities afforded by generative AI for business, and how those technologies could be adopted in novel ways in a business context. Requiring students to use AI tools to create multiple

distinct types of output was thus designed to move students beyond using AI to generate text, and towards innovative, creative uses of AI: using AI as a co-creator in their future productivity (Mollick, 2024) rather than as a shortcut in cognitive processes.

This assessment was more successful than anticipated. Students engaged strongly with the creativity afforded by the assessment, with students developing working chatbots for businesses, social media marketing campaigns, entire websites and functioning mobile applications – and in their reflections, many students commented that they were surprised by what they had accomplished. This assessment task thus moved assessment into being a form of learning rather than just a tool to assess learning (Earl, 2006), and successfully encouraged students to develop their skills beyond what was taught in the classrooms. Some students also went on to use this assessment as a portfolio in job applications, making it relevant to their future career as well as their studies.

In addition to this portfolio, the second assessment for the module was designed as a ‘Professional Discussion’. In a digital module, it was felt important to not lose the focus on criticality and reflexivity, and so this assessment was entirely in person with students required to take part in an interview with the module leader, where they would be given a case study of business AI use they had not seen before and had to discuss the case study, relating it to their wider reading and knowledge of AI.

This assessment served two functions. Firstly, students were encouraged to use AI tools to help in their preparation for the task – such as in identifying sources to research, summarising information and assimilating ideas quickly – thus building their AI literacy and again encouraging a more critical use of AI. Secondly, by holding the assessment in person and encouraging students to give their own opinion on a case study they hadn’t previously seen, students developed their criticality and confidence: skills that match both the subject benchmark statement for Business and Management degrees (QAA, nd) and the demands of employers (CMI, 2021).

This professional discussion assessment was very successful, both in terms of students’ engagement with the task and also in being an authentic and valid measure of each student’s reflective ability and critical thought around AI. The one downside of this assessment was in its time-intensive nature: each discussion took almost 30 minutes, as well as additional time for marking and moderation, and this was a large addition to the academic workload to undertake and mark. Nonetheless, it was a rewarding experience to see students engaging so thoughtfully with the module content through the assessment, and this is a format we intend to continue to roll out moving forward – not only in this AI-based module but across a spectrum of modules on the course.

Outcomes

The module has now been run twice, with development to the module each year. Feedback on both the module and the assessment mix from students has been largely positive, with students particularly expressing positive feedback on the portfolio assessment, as one student from the 2024-25 cohort expressed:

“It’s honestly one of the most valuable things I’ve studied at university. One of the best things about the module for me was about how varied and interactive the lessons and assessments were. As someone who is really hands on and practical, I think that was great that we were actually physically using tools. I love that we actually physically went in and were able to create a portfolio ourselves as one of the assignments to showcase what we have learned. I know some people see AI as a cheat code or AI not requiring you to use your brains as much, but I think more modules should have AI introduced into it. I think this module had the right balance of theory and practice.”

(Student, 2024-25 cohort)

It will be important for this module and assessment to further evolve year on year. The biggest challenge so far has been the rapid development of AI technologies: what was difficult and innovative for students in 2024-25 was commonplace in 2025-26, and so assessment rubrics linked to innovation and creativity required significant standardisation for the second run in order to ensure the validity of student grades. Much of the module content also required significant updates in the second year to keep up to date with the technological changes.

While there has undoubtedly been a significant time commitment in developing and maintaining this module – especially given the need to rewrite large parts of the module each year to keep up with advancements in the technology – the feedback from students and the growth in their AI literacy illustrates the importance of this sort of module and assessment within the field of Business and Management. Anecdotal evidence from the course team and students following the module over the last year shows that it has led to an increase in thoughtful and critical use of AI by students in their other assessments and modules, along with a decrease in misconduct issues due to inappropriate use of AI. Qualitative research is planned focusing on the current cohort of students to examine this further.

The key lesson learnt has been that explicitly teaching AI literacy has moved AI from a ‘taboo’ topic or a form of misconduct to a key graduate attribute for students, giving them a key skill for employability as well as encouraging criticality and reflexivity in their use of AI. Given the growth in AI and the challenges it poses to traditional assessment formats in many disciplines, this lesson may well be transferable to other disciplines and higher education contexts.

References

- Arowosegbe, A, Alqahtani, J S and Oyelade, T (2024) 'Perception of generative AI use in UK higher education', *Frontiers in Education*, 8. Available at: doi.org/10.3389/feduc.2024.1463208
- Cheng, M, Adekola, O, Albia, J and Cai, S (2022) 'Employability in higher education: a review of key stakeholders' perspectives', *Higher Education Evaluation and Development*, 16 (1): 16-31. Available at: doi.org/10.1108/HEED-03-2021-0025
- CMI (2021) 'Work ready graduates: building employability skills for a hybrid world'. Available at: www.managers.org.uk/knowledge-and-insights/research/building-employability-skills/ [accessed 1 February 2026].
- Kennedy, E (2024) 'Teaching AI in business schools: a pathway to future-ready graduates', *Journal of Instructional Pedagogies*, 30. Available at: eric.ed.gov/?id=EJ1459914
- Mollick, E (2024) *Co-Intelligence: living and working with AI*. London: W H Allen.
- Portocarrero Ramos, H C, Cruz Caro, O, Sánchez Bardales, E, Quiñones Huatangari, L, Campos Trigoso, J A, Maicelo Guevara, J L and Chávez Santos, R (2025) 'Artificial intelligence skills and their impact on the employability of university graduates', *Frontiers in Artificial Intelligence*, 8. Available at: doi.org/10.3389/frai.2025.1629320
- QAA (nd) 'Subject Benchmark Statement: Business and Management'. Available at: www.qaa.ac.uk/the-quality-code/subject-benchmark-statements [accessed 1 February 2026].

Embedding work-integrated learning through human-centred design for an AI-driven environment

Nina Edward, Associate Professor in Design; Wai Hoe Yan, Casual Academic and Yizhi Wu, Casual Academic, University of Sydney, Australia

Background

IDEA9202 Experience Design Studio is a core 12-credit postgraduate studio unit within the Master of Interaction Design and Electronic Arts (MIDEA) at the University of Sydney. The cohort is predominantly international students with design backgrounds, alongside students from Information Technology, Business and Communication. This disciplinary diversity makes the unit a critical site for developing shared design capabilities, professional identity and industry readiness at the postgraduate level.

This work addresses a misalignment between prevailing higher education assessment practices and the realities of contemporary design practice, particularly amid increasing AI integration. Assessment remains dominated by tasks privileging content recall, standardisation and product-focused outputs – forms increasingly susceptible to automation by generative AI (Kizilcec et al, 2024). In design education, this risks narrowing learning to artefact production while undermining empathy, ethical judgement, collaboration, creative problem-solving and independent professional judgement (Gerlich, 2025; Zhai et al, 2024).

In response, this case study presents a redesign of assessment in IDEA9202 that foregrounds process-oriented, human-centred and industry-facing learning outcomes robust in AI-rich environments. The unit develops students' capacity to apply research-led interaction design processes; generate, evaluate and refine novel design concepts; prototype interactive systems; attend to aesthetic and experiential qualities; communicate persuasively to specialist and non-specialist audiences; collaborate effectively in teams; and work independently through reflection, self-evaluation and incorporation of feedback. Ethical and critical use of digital tools, including AI, is explicitly embedded within project work.

The assessment redesign is guided by three objectives:

- + embedding authentic work-integrated learning (WIL) through sustained client engagement
- + cultivating distinctly human capabilities alongside critical digital fluency
- + aligning assessment with sector expectations around graduate employability, academic integrity and future-ready education.

The work is informed by an integrated pedagogical framework. Human-Centred Design (HCD) provides the primary foundation, foregrounding empathy, iteration, reflexivity and co-creation as core competencies. Authentic assessment theory informs task design, emphasising performances that mirror professional practice and require judgement in complex contexts (Kember and McNaught, 2007).

WIL operates as the structural framework for the assessment redesign. Consistent with the Higher Education Standards Framework (Threshold Standards) 2021, WIL is defined as learning undertaken in a work context as part of course requirements, including industry-partnered studio projects (TEQSA, 2022). WIL is widely recognised for supporting employability by enabling students to apply theory in authentic contexts, develop professional networks and prepare for workforce transition (Patrick, Peach and Pocknee, 2009; University of Sydney, 2025).

In IDEA9202, this approach is enacted through a real client project with the Multi-Arts Pavilion, mima (MAP mima), located in Lake Macquarie, New South Wales. MAP mima is a contemporary museum and digital art venue that hosts immersive and interactive installations, providing an authentic professional context for students to engage with a client, respond to real briefs and negotiate the complexities of interaction and experience design practice. Disciplinarily, the redesign responds to long-standing challenges in interaction and experience design education, where professional competence is demonstrated not only through final artefacts, but through iterative processes, ethical decision-making, teamwork and sustained stakeholder engagement.

Approach

Over the past two years and across two semesters, three core pedagogical approaches were employed to support student learning within this unit: work-integrated learning (WIL), the embedding of Human-Centred Design (HCD) methods, and the integration of digital tools alongside design methodologies and soft-skills. Collectively, these strategies cultivated a learning environment that closely aligns with contemporary professional practice and prepares students for future trajectories in the design field.

Authentic assessment through WIL with MAP mima

The first approach involved the adoption of WIL as a mechanism for constructing an authentically situated assessment framework. Students collaborated directly with MAP mima and were required to engage with genuine design briefs rather than abstract or hypothetical scenarios. This arrangement exposed students to the operational complexities of professional design work, including financial constraints and the imperative to generate outcomes that are simultaneously creative and practically achievable.

Working with external clients required students to negotiate project intentions, articulate design decisions with clarity and integrate stakeholder perspectives into iterative design revisions. Students also practised translating conceptual design intentions into coded artefacts to enable more efficient project handovers – an increasingly consequential skill in contemporary design workflows. Principles of accessibility and inclusivity were embedded throughout, reinforcing the ethical commitments expected of emerging designers. Together, these experiences were intended to strengthen students' professional readiness and facilitate a more confident transition from academic learning into industry practice.

Embedding HCD across the design process

The second approach centred on embedding HCD across all phases of project development. The design process commenced with in-depth site research, during which students conducted detailed investigations of MAP mima and visited analogous locations to expand their contextual understanding. HCD methods supported students in developing an informed sensitivity to local spatial, cultural and environmental conditions. The methods encouraged multi-perspectival interpretation of place and ensured that students' early conceptual thinking was grounded in contextual specificity.

The subsequent phase focused on user understanding. HCD techniques enhanced students' observational acuity, analytical proficiency, communication skills and ethical awareness. Treating users as active participants in the design process enabled students to identify needs, expectations and potential friction points embedded in their emerging design responses.

The delivery phase culminated in the production of prototypes that were presented at a testing fair attended by clients and industry partners. Students were required to justify their approach, address feedback and outline improvements. Through both digital and hands-on prototyping, students gained experience with iterative design workflows that reflect established industry practice.

Integrating digital tools, design methods and soft skills

The third approach involved an integrative pedagogical model that positioned digital tools, design methods and soft skills as interdependent components of the learning experience. Digital platforms such as Figma provided a collaborative environment where students could organise their work in real time. As students increasingly incorporated AI-assisted tools into their workflows, the teaching team emphasised that such technologies enhance design execution but do not replace thoughtful, human-centred decision-making.

Design methods continued to anchor the unit, guiding students through research, ideation, testing and implementation. Their consistent use ensured that technological tools were deployed with intentionality and methodological rigour.

Soft skills were positioned as indispensable attributes of competent designers. Despite rapid technological advancement, it remains critical to reaffirm that design is, at its core, a human-centred discipline. Framing design around human value and criticality encouraged students to move fluidly between creative practice, digital augmentation and ethical reflection.

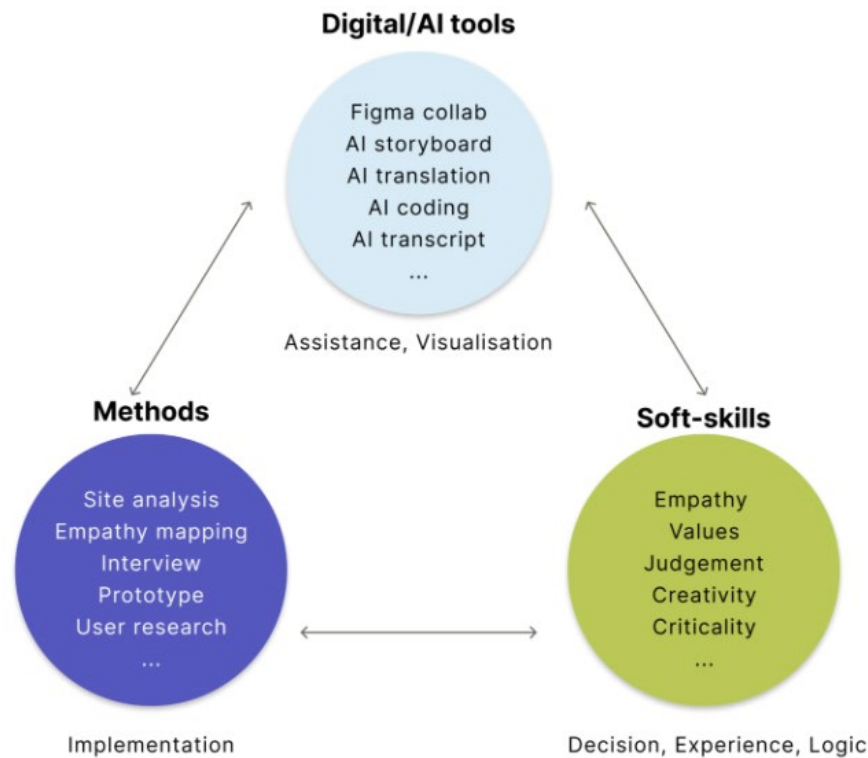


Figure 1. Integrative model of AI tools, methods and soft skills

In combination, these three approaches created a cohesive pedagogical framework that prepared students for real-world design practice. This unit supported students in developing confidence within an AI-enhanced design landscape while maintaining the empathy and criticality that lie at the core of design disciplines.

Outcomes

This case study compared student learning experiences under two assessment models in IDEA9202; a conceptual brief in Semester 2, 2024; and an authentic WIL brief in Semester 1, 2025. Analysis was based on University of Sydney Unit of Study Survey (USS) data from both cohorts.

In 2024, the unit enrolled 189 students with 84 survey responses ($n = 84$), representing a 44 percent response rate. In 2025, enrolment was 137 students with 46 responses ($n = 46$), representing a 34% response rate. The 10% difference in response rates required statistical adjustment to ensure fair comparison.

Four key metrics; drawn from 16 USS survey items; were selected for analysis because they most closely reflected the aims of the assessment redesign: human-centred capability, professional relevance and digital literacy. The four metrics are: I developed my ability to think independently about problems; I developed my capacity to respond constructively to challenge; I developed the ability to practically apply knowledge of the field(s) I am studying; I developed my ability to work effectively with digital and online tools and information.

Comparative analysis used mean scores, weighted means and Welch t-tests to account for unequal sample sizes and confirm statistical reliability.

Human-centred design capabilities

The first set of measures focused on human-centred skills, specifically independent thinking and constructive response to challenge.

For the independent thinking item, the 2024 cohort (n = 84) recorded a mean score of 4.16. In 2025, this increased to an adjusted mean of 4.29 (n = 46). For the second human-centred item measuring constructive response to challenge, the mean score rose from 4.14 in 2024 (n = 84) to 4.23 in 2025 (n = 46).

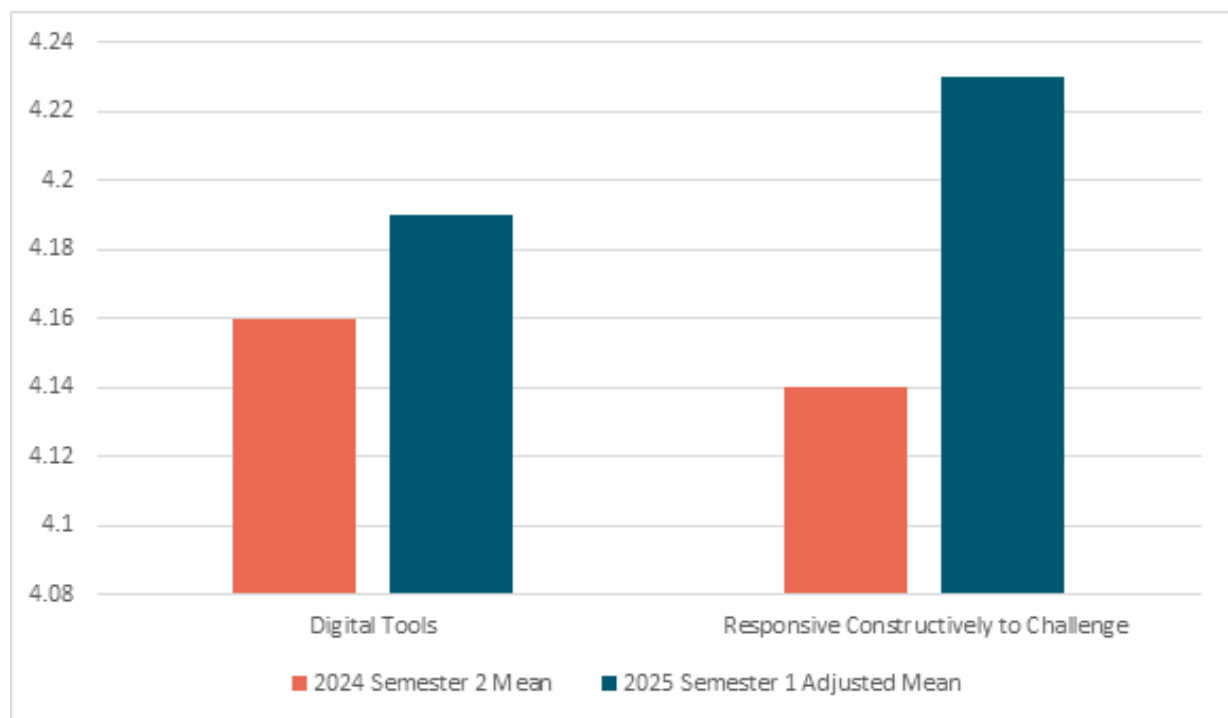


Figure 2: Comparison of human-centred skill indicators with mean scores and counts.

These improvements demonstrate a consistent upward trend in students' perception of their capacity to think critically and respond productively to complex design problems. Qualitative feedback supported this pattern, with students describing the unit as challenging but enjoyable and emphasising the value of hands-on, socially engaged learning activities.

Authentic application and professional relevance

The largest improvement was observed in the item measuring perceived professional relevance and real-world application. In 2024, this item recorded a mean score of 3.89 (n = 84). In 2025, the raw mean increased to 4.22 (n = 46). After adjusting for differences in response rates, the weighted mean for 2025 was calculated as 4.01, with a final adjusted mean of 4.19 following Welch t-test correction. This represents an approximate 4.5% improvement from the weighted baseline of 4.01 to the adjusted value of 4.19, achieved at a 90% confidence level. These findings indicate a meaningful enhancement in students' perception of their ability to apply knowledge in authentic, real-world professional contexts.



Figure 3. Improvement in professional relevance indicator.

Student comments consistently emphasised the importance of investigating clients and sites and understanding how real cases are designed. These responses indicate that interaction with live clients enabled students to connect academic learning with professional practice, strengthening both understanding and motivation.

Digital literacy and adaptability

Measures of digital fluency also showed positive movement. The digital literacy item recorded a mean score of approximately 4.08 in 2024 (n = 84), rising to an adjusted mean of 4.21 in 2025 (n = 46). Students reported that collaborative learning and experimentation with diverse digital tools enhanced their confidence and ability to communicate design ideas effectively.

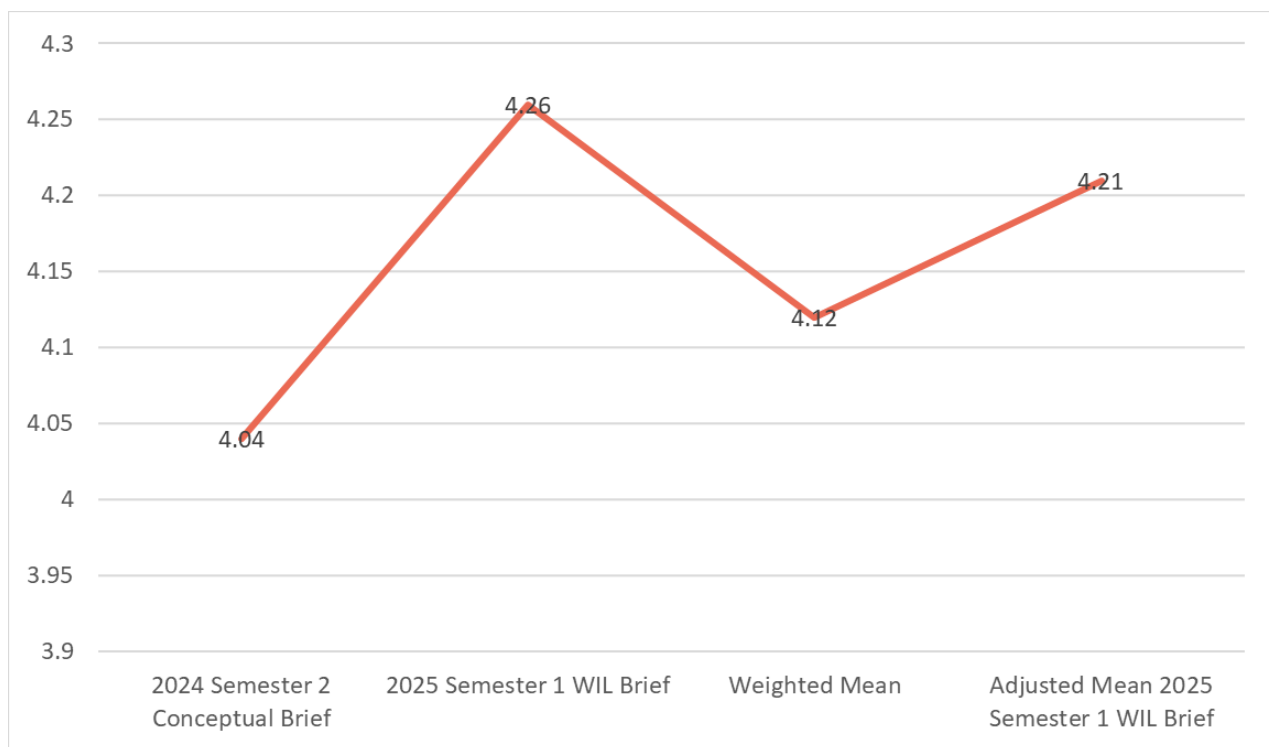


Figure 4. Comparison of digital literacy outcomes across cohorts.

This environment promoted peer-supported learning and practical problem solving rather than tool-driven task completion. Students learned to select digital methods strategically, adapt quickly to new technologies and communicate ideas more effectively through diverse formats. These outcomes indicate growth not only in technical competence but also in critical digital judgement and adaptability; skills essential for professional practice in an AI-augmented design landscape.

Key insights and implications

Across the four analysed metrics, results consistently favoured the authentic assessment cohort. Students showed stronger engagement when learning involved challenge, collaboration and authentic application. The inclusion of live, client-based projects strengthened their ability to apply disciplinary knowledge in meaningful contexts. Integration of diverse digital platforms enhanced creative confidence, methodological flexibility and the capacity to learn new technologies independently.

These outcomes suggest that authentic assessment better supports the development of complex human capabilities that are difficult to automate, including critical thinking, empathy, teamwork and ethical judgement, while simultaneously strengthening digital fluency.

Transferability and future directions

Further development is needed to consolidate and extend these outcomes. Future work should include collecting data across additional cohorts to support longitudinal analysis, incorporating complementary indicators such as client feedback, and refining assessment rubrics to more explicitly recognise process skills and responsible AI use. Expanding industry partnerships and examining which components of the WIL and HCD approach most strongly influence learning will further support ongoing improvement.

This case study raises a broader question for contemporary education: how can institutions balance AI-driven efficiency with the cultivation of human creativity, empathy and ethical judgement? Findings from IDEA9202 suggest that embedding HCD and WIL within assessment offers a constructive pathway forward. Further research may explore how this model can be adapted across disciplines, professional contexts and delivery modes to support more resilient, AI-ready assessment practices.

References

Gerlich, M (2025) 'AI tools in society: Impacts on cognitive offloading and the future of critical thinking', *Societies*, 15 (1): 6. Available at: doi.org/10.3390/soc15010006

Kember, D and McNaught, C (2007) *Enhancing university teaching*. Abingdon: Routledge.

Kizilcec, R F, Huber, E, Papanastasiou, E C, Cram, A, Makridis, C A, Smolansky, A, Zeivots, S and Radulescu, C (2024) 'Perceived impact of generative AI on assessments: Comparing educator and student perspectives in Australia, Cyprus, and the United States', *Computers and Education: Artificial Intelligence*, 7: 100269. Available at: doi.org/10.1016/j.caeai.2024.100269

Patrick, C, Peach, D and Pocknee, C (2009) *The WIL (Work Integrated Learning) report: a national scoping study*. Australian Learning and Teaching Council.

Tertiary Education Quality and Standards Agency (2022) *Guidance note: Work-integrated learning*. Version 2.0. Available at: <https://www.teqsa.gov.au/sites/default/files/guidance-note-work-integrated-learning-v2-0.docx>

Zhai, C, Wibowo, S and Li, L D (2024) 'The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review', *Smart Learning Environments*, 11 (1): 28. Available at: doi.org/10.1186/s40561-024-00316-7

Flawed by design: embedding AI to drive authentic assessment and integrity

Professor Rania Shibl, Southern Cross University, Australia, and Dr Judy Watson, University of the Sunshine Coast, Australia

Background

The rapid and widespread adoption of generative artificial intelligence (AI) tools in higher education has intensified long-standing debates about academic integrity, authenticity and the purpose of assessment. Tools such as ChatGPT and other large language models are now capable of producing syntactically correct code, coherent explanations and complete solutions to many assessment tasks traditionally used in computing and data-related disciplines. While these technologies offer significant opportunities to support learning, they also challenge educators to reconsider what knowledge, skills and attributes assessments are designed to measure.

In disciplines such as information and communications technology (ICT), the impact of generative AI is particularly acute. Tasks that previously served as reliable indicators of individual competence, such as writing SQL queries or designing database schemas, can now be completed, at least superficially, through automated means. This raises critical questions: if AI tools can generate code, what is the educational value of assessing code production alone? How can assessment practices evolve to ensure they continue to measure student understanding, judgement and accountability in an AI-enabled world?

This case study is situated within a second-year undergraduate course in database design (ICT211) at the University of the Sunshine Coast (UniSC), enrolling approximately 200 students. The course focuses on foundational concepts in relational database design, including schema development, data integrity constraints and SQL implementation. Historically, assessment in this course followed a conventional model, requiring students to design and implement database structures based on a provided case study. However, the increasing accessibility of generative AI tools significantly undermined the reliability of these tasks as indicators of genuine learning.

The central challenge addressed in this work was the design of a summative assessment that remained authentic, academically rigorous and scalable, while acknowledging rather than prohibiting the presence of AI tools. Rather than attempting to restrict or detect AI use, the assessment sought to reposition AI as part of the learning environment and to shift the focus of assessment from content generation to content evaluation and responsibility.

The work is informed by several complementary theoretical and conceptual frameworks that collectively support a shift toward authentic, applied and AI-aware assessment design. Authentic assessment theory argues that students should demonstrate learning through tasks that mirror real-world professional practice rather than through decontextualised or purely abstract exercises. Wiggins' (1990) seminal work on authentic assessment

emphasised that learners should engage in the same kinds of complex tasks faced by practitioners, thereby bridging the gap between education and practice. In higher education contexts, authentic assessment has been shown to enhance engagement, motivation and the development of transferable, 21st-century skills by requiring students to integrate knowledge and apply it in meaningful situations (Vlachopoulos and Makri, 2024). In the context of database design, this means evaluating not only whether students can write syntactically correct SQL, but whether they can diagnose errors, interpret requirements and implement solutions that would function in a real organisational setting.

Constructivist learning theory further underpins this approach. From a constructivist perspective, learners actively construct understanding through problem-solving, interpretation, dialogue and reflection rather than passively absorbing information (Tam, 2000). Foundational theorists such as Piaget and Vygotsky highlighted the active and social nature of learning, arguing that knowledge is built as students grapple with authentic problems and reflect on their experiences. In practical terms, constructivist pedagogy privileges inquiry-based and open-ended tasks that require learners to test ideas, receive feedback and refine their understanding (Chand, 2024). The use of deliberately flawed scripts in this case study aligns closely with this paradigm: students must interpret the problem, diagnose inconsistencies and iteratively refine their solutions, thereby constructing knowledge through engagement with complexity rather than reproducing memorised procedures.

Recent scholarship on AI in education further reframes assessment design in light of generative technologies. Rather than positioning AI solely as a threat to academic integrity, emerging research advocates treating AI as a cognitive partner within the learning process (Elshall and Badir, 2025). This perspective calls for assessments that make student thinking visible and foreground distinctly human capabilities such as critical judgement, ethical reasoning, metacognitive oversight and accountability. Since generative AI can now produce plausible code or written responses, sustainable assessment in the AI era must move beyond evaluating surface-level outputs and instead examine how students critique, validate, adapt and take responsibility for AI-assisted work. Sector-wide discussions and contemporary research emphasise that assessment redesign should leverage AI for what it does well while ensuring that human expertise remains central (Corbin, Dawson and Liu, 2025; Khlaif et al, 2025). This case study aligns with these calls by embedding AI within the assessment process in a constrained and purposeful manner, shifting the focus from content generation to evaluation, professional judgement and accountability in an AI-enabled world.

Approach

To address these challenges, a constrained AI-supported assessment model was developed and implemented. Central to this approach was the creation of 'Liam Chat', a simulated AI client designed to generate assessment inputs rather than assessment outputs. Liam Chat was built using OpenAI's API and deployed through a ShinyApps.io

interface, providing a lightweight, web-based tool accessible to students during the assessment period.

The assessment task required students to review, correct and complete a flawed SQL CREATE TABLE script and to populate the database with appropriate sample data based on a fictional business case, *The Smart Café*. Rather than providing all students with the same script and dataset, Liam Chat generated unique, personalised inputs for each student. These inputs were deliberately flawed, containing errors in data types, missing or incorrect constraints, or logical inconsistencies that required diagnosis and correction.

A key design decision was to tightly constrain Liam Chat's capabilities. Although initially envisioned as a tool for interactive requirement gathering, its functionality in this iteration was deliberately limited to two outputs only: (1) a flawed SQL schema and (2) a dynamically generated dataset. Liam Chat did not provide explanations, hints, or corrective feedback. This constraint was essential to maintain alignment with learning outcomes and to prevent the AI from inadvertently completing the task on behalf of students.

Students were explicitly informed that the scripts generated by Liam Chat would be incorrect, but that the nature and location of the errors would vary. This uncertainty mirrored real-world professional contexts, where system specifications are often incomplete or inaccurate and where practitioners must exercise judgement rather than follow instructions blindly. Importantly, students were not prohibited from using external generative AI tools; instead, the task was designed such that superficial reliance on AI-generated solutions would be ineffective unless students could evaluate and adapt those solutions appropriately.

The assessment was delivered over a four-week period. Students accessed Liam Chat to generate their individual scripts and datasets, worked independently to correct and implement the database, and submitted their final SQL code via the learning management system. Marking criteria were deliberately simplified and aligned with professional standards: whether the code executed without errors, whether the data was correctly inserted, and whether the final implementation adhered to the business rules specified in the case study.

Staff involvement focused primarily on assessment design, tool configuration and quality assurance rather than ongoing intervention during the assessment period. Once deployed, the automated generation of scripts and data significantly reduced administrative overhead and ensured consistency in assessment difficulty across the cohort, despite the individualisation of inputs.

In terms of implementation, the development of Liam Chat can be achieved through two primary approaches, each with different trade-offs. The approach adopted in this case used OpenAI's API integrated within a ShinyApps.io interface, allowing students to access the tool without needing individual accounts. This supported equitable access but required initial technical expertise, ongoing maintenance and minimal costs associated with API token usage. The initial development was completed over approximately two to three weeks, equating to around 20 to 30 hours of focused development time, including prompt design,

interface setup, testing and refinement to ensure stable and constrained outputs. An alternative approach is to create a custom GPT within ChatGPT, which removes the need for deployment infrastructure and significantly reduces maintenance overhead. However, this requires students to have access to a (free) ChatGPT account. While this may raise initial access considerations, the increasing prevalence of student use of such tools suggests this is becoming less of a barrier. In both approaches, reliable internet access remains a prerequisite. These options provide flexible pathways for implementation depending on institutional context, technical capability and equity considerations.

Outcomes

The implementation of the constrained AI-supported assessment yielded several important outcomes for both students and staff.

Student outcomes

The assessment revealed substantial variation in student understanding and capability. Students who possessed a strong grasp of database design principles were able to diagnose errors, adapt to the flawed inputs and produce valid, functional database implementations. In contrast, many students who relied heavily on generative AI tools without sufficient conceptual understanding struggled to complete the task successfully. This outcome was notable, as AI use was not restricted; rather, the task design itself exposed the limitations of uncritical AI reliance.

The authenticity of the task played a significant role in these outcomes. Students were required to take responsibility for the correctness and consequences of their implementations, reflecting the realities of professional practice. While this resulted in higher failure rates than in previous iterations of the course, it also provided a more accurate representation of student readiness and highlighted areas requiring further instructional support. Informal observations and student discussions indicated that many found the task more challenging than previous assessments, particularly due to the variability and lack of a single “correct” solution path. However, students who engaged deeply with the task reported that it more closely resembled real-world problem-solving and highlighted the importance of understanding, rather than reproducing, SQL logic.

Staff outcomes

From a teaching perspective, the assessment significantly simplified marking and moderation processes. The focus on executable code and adherence to business rules reduced subjectivity and ambiguity in grading decisions. The personalised nature of the assessment inputs effectively eliminated opportunities for collusion, and no cases of plagiarism or academic misconduct were reported.

The model proved highly scalable, accommodating a large cohort without a proportional increase in staff workload. Once established, the assessment design was reusable and adaptable, offering long-term sustainability benefits.

Key messages and transferability

This case study demonstrates that constrained AI can be used to enhance, rather than undermine, assessment authenticity and integrity. By shifting the focus from production to evaluation, educators can design assessments that focus on human judgement, accountability and professional responsibility.

The approach is transferable to other disciplines and levels, particularly where learning outcomes involve interpreting, validating, or correcting complex artefacts. Potential applications include data analytics, programming, systems analysis, business rule modelling, and even non-technical fields where students must critique AI-generated content.

Future development

Future iterations of this work will explore extending Liam Chat's functionality to support requirement gathering through natural language interaction, enabling students to ask clarifying questions and experience simulated client interviews. Further research will also examine how scaffolded AI-supported assessments can be integrated across programmes to progressively develop students' AI literacy, ethical awareness and professional judgement.

References

- Chand, S P (2023) 'Constructivism in education: Exploring the contributions of Piaget, Vygotsky and Bruner', *International Journal of Science and Research*, 12 (7): 274-278.
- Corbin, T, Dawson, P and Liu, D (2025) 'Talk is cheap: why structural assessment changes are needed for a time of GenAI assessment and evaluation in higher education', *Assessment & Evaluation in Higher Education*, 50 (7): 1087-1097.
- Elshall, A S and Badir, A (2025) 'Balancing AI-assisted learning and traditional assessment: the FACT assessment in environmental data science education', *Frontiers in Education*, 10: 1596462.
- Khlaif, Z N, Alkouk, W A, Salama, N and Abu Eideh, B (2025) 'Redesigning assessments for AI-enhanced learning: a framework for educators in the generative AI era', *Education Sciences*, 15 (2): 174.
- Tam, M (2000) 'Constructivism, instructional design, and technology: Implications for transforming distance learning', *Educational Technology and Society*, 3 (2): 50-60.
- Vlachopoulos, D and Makri, A (2024) 'A systematic literature review on authentic assessment in higher education: best practices for the development of 21st century skills, and policy considerations', *Studies in Educational Evaluation*, 83: 101425.
- Wiggins, G (1990) 'The case for authentic assessment', *Practical Assessment, Research and Evaluation*, 2 (2).
-

When the threat becomes a tool: using generative AI to create authentic assessments for final year Life Sciences students

Dr Emma Reid, Lecturer, University of Glasgow, UK

Background

The use of generative AI (GenAI) tools by undergraduate students to complete in-course assessments has become an increasing concern to teaching staff within academic institutions, with many educators viewing AI platforms as a threat to the academic integrity of authentic assessment (Williams, 2025). It also raises concerns about academic competence, and development of cognitive skills (Ipek et al, 2023), where the use of GenAI tools may limit student engagement with learning activities, which aim to increase depth of knowledge and critical thinking skills. As a result, HE institutions are now reflecting on how to develop policy relating to acceptable use of GenAI within the educational framework and considering how to adapt pedagogical practice accordingly. Educators are beginning to realise that GenAI can be successfully incorporated into course content to enhance student learning and increase engagement. Further, it can also effectively reduce the time spent by academics on writing tasks (Noy and Zhang, 2023), allowing educators to redirect their focus to student-facing activities.

Traditional assessment styles, such as essays, MCQs etc, are becoming increasingly vulnerable to the use of GenAI, reducing their academic value (Owen and Kay, 2025). Many educators are shifting assessment styles towards the use of performance-based activities, including oral presentations, collaborative assignments and practical evaluations (Owen and Kay, 2025). However, the rapid advancements in GenAI does not mean that we need to dismiss traditional assessments completely. Current discussion focuses on the use of GenAI in the creation of authentic assessments. My interest centres on the incorporation of GenAI into traditional in-person examinations. End-of-course examinations can result in students experiencing test anxiety and assessment related stress (Putwain, 2007) and many fear 'the blank page' in exams, which can have a detrimental impact on academic performance (von der Embse et al, 2018).

However, redesigning assessments can successfully alleviate test anxiety, while still effectively evaluating the same learning outcomes (Hurren, Rutledge and Garvin, 2006). Student critique of GenAI outputs is an approach which has been recently used within Life Sciences courses at the University of Glasgow in earlier years of study, but it has only been used within in-course assessment, not within in-person examinations. The use of GenAI critique activities within coursework assignments is susceptible to inappropriate student use of GenAI, since GenAI now has the capability to critique its own outputs. Asking students to critique GenAI outputs under examination conditions eliminates this risk.

To this end, I aimed to undertake a pilot case study to use GenAI tools to create an authentic assessment for the newly formed Physiological Disorders: Mechanisms and Treatment Challenges module. The assessment would take the form of an end-of-course examination to effectively evaluate the intended learning outcomes of the course and simultaneously alleviate test anxiety in the student cohort.

Approach

I designed a new course called Physiological Disorders: Mechanisms and Treatment Challenges, as an elective module for final year Human Biology and Physiology undergraduates within the University of Glasgow. The course assessment profile comprises a combination of in-course assessment: a flash talk (20% of final course grade) and 2,000-word literature review (30% weighting), and an end-of course, in-person examination (50% of final course grade). I used GenAI to create the end of course examination paper, where students would be asked to critically evaluate a GenAI essay on a core course topic, in-person, under exam conditions. Using Microsoft Co-pilot, which is a GenAI tool approved by the University of Glasgow, I generated an 800-word essay, using the prompt:

‘Produce an 800-word scientific overview of diabetic retinopathy’.

Note that the topic of diabetic retinopathy is an example of one of the physiological disorders covered in the course and could be substituted within the prompt for any other relevant course topic.

After fact checking the essay to ensure that the themes covered in the GenAI output linked to the content covered within the relevant teaching session, I then introduced two factual inaccuracies into the essay content. One error was designed to be obvious, and could be easily detected, and the other error was more subtle, and required greater depth of understanding of the topic to identify it.

Students were asked to critique the essay in 90 minutes, under exam conditions, with the following prompt questions provided within the exam paper, to guide their evaluation:

- 1 Is the essay factually correct? Are there any inaccuracies in the content?
- 2 What additional detail would you include in each section of the essay to make it more informative for the scientific community?
- 3 Would you add any additional information that hasn't been addressed in the essay? If so, what would you include?

In advance of the examination, students were advised on the format of the exam and the prompt questions but were not provided with any information relating to the specific topic or given access to the AI-generated essay prior to the exam. This was a crucial component to secure the integrity of the assessment, due to GenAI's capability to critique its own outputs. Students were also encouraged to spend approximately 30 minutes reading the AI essay and use the remaining hour of the exam to write their critical evaluation. Exam scripts were assessed independently by two markers, who agreed an overall grade and reconciled feedback comments for each student's submission.

Outcomes

The assessment was highly successful and all students in the cohort passed the exam ($n=23$), with over 60% of students achieving an A or B grade. The University of Glasgow uses a 22-point grading scale, where each grade has a numerical value: 22-18 is equivalent to a grade of A1 to A5, 17-15 equates to grades B1-B3 etc. A grade of D3, or 9, on the numerical scale, is the pass threshold.

The mean class grade for the examination was a B2 (16) grade, which is directly comparable to the mean grades achieved for both in-course assignments, where the mean grades awarded for the flash talk and the literature review, were B1 (17) and B2 (16), respectively (One-way ANOVA, $P=0.27$, Figure 1). This demonstrates consistency across the course assessment profile in evaluating the range of intended learning outcomes of the course, using varied styles of assessment.

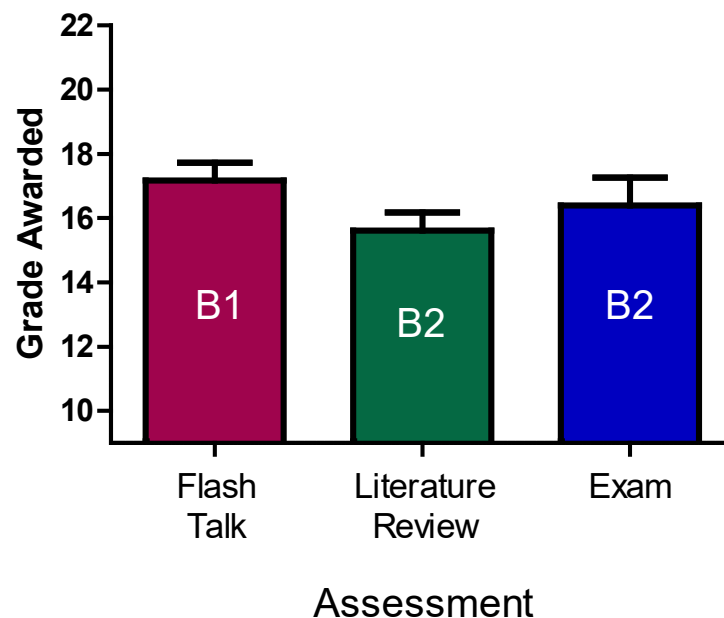


Figure 1. Grade profile of course assessments. Data presented as mean $\pm$ SEM. Y-axis shows numerical 22-point grading scale, with mean equivalent grade for each assessment shown in white. One-way ANOVA, $P=0.27$ (ns).

Within Life Sciences at the University of Glasgow, we typically observe that students achieve lower grades for in-person examinations, when compared to in-course assessment. However, this was not the case in the present study, where grades achieved by individual students were consistent across the course assessment profile, with no significant difference in the grades achieved for the flash talk, literature review and the end of course examination (Repeated Measures ANOVA, $p=0.1$, Figure 2).

The GenAI assessment shifted the examination focus from the traditional memorisation approach to positioning students as critical evaluators, with the emphasis on analysis and

application of knowledge. It also enabled students to recognise the limitations of generative AI by encouraging them to actively seek out misinformation, or poorly explained concepts. Students were also able to demonstrate their depth of understanding by outlining key points in greater detail than that presented within the 800-word limit of the AI-generated summary. Several students also verbally reported to the course coordinators that they felt less anxious in the period leading up to the examination, and were confident that they could perform well in the assessment. Both the format of the exam, and the advance knowledge of the prompt questions, contributed to student wellbeing by reducing the 'fear of the blank page'.

From my own personal perspective, the process of marking this assessment was an enjoyable experience, which provided great insight into the depth of student understanding of the course content. In particular, student enthusiasm for the topic was clearly demonstrated in their critiques. This is an element which is often lacking in more traditional essay type assessments, where content is simply presented in a descriptive style, without evidence of engagement.

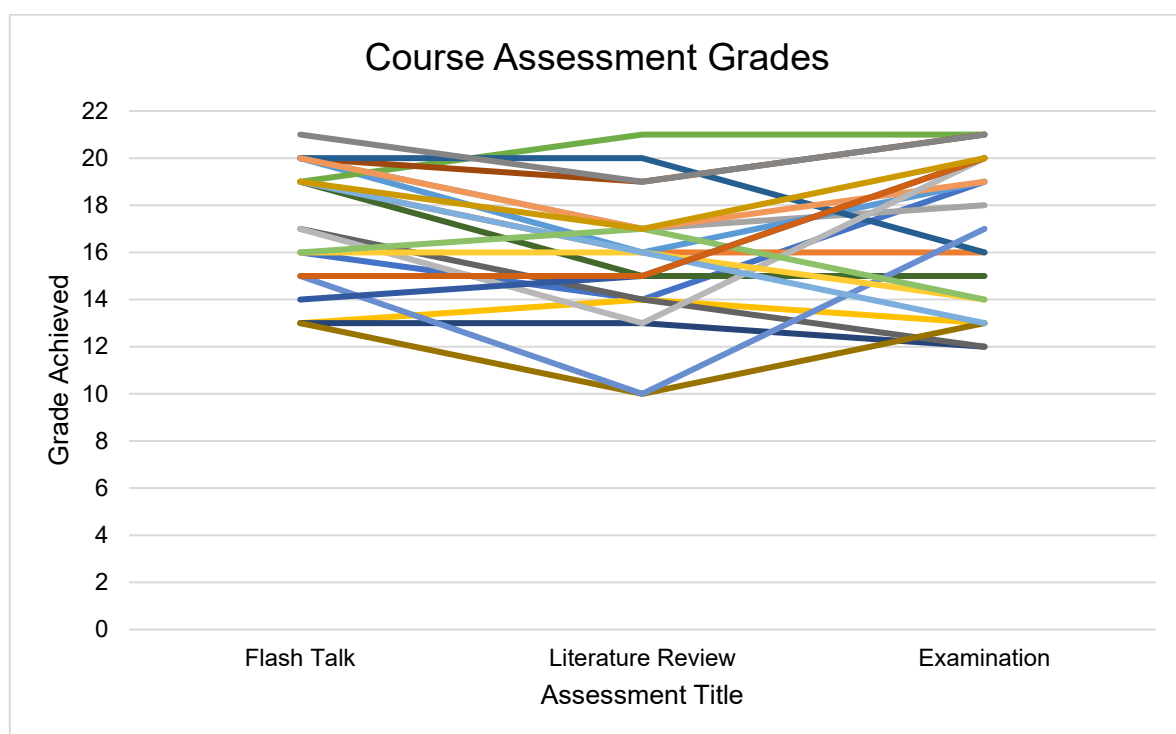


Figure 2. Student performance across the course assessment profile. Data presented shows the individual grades awarded for each assessment for all students within the cohort.

Repeated Measures ANOVA, $P=0.1$ (ns)

However, it is important to recognise the limitations of the current approach. This case study only evaluated a single small cohort of students, and it will be necessary to assess course outcomes over several years to investigate the long-term value of this assessment approach. Additionally, the current design of the GenAI assessment within this case study evaluates knowledge and understanding of a single course topic. The assessment could be

amended to evaluate student understanding from a range of topics or provide a choice of course topic within the examination. This has yet to be explored, but merits consideration in the development of future assessments. Further, the use of GenAI-based assessments could be evaluated in a wider context within Life Sciences. This approach is currently being considered within several Life Sciences degree programmes within the University of Glasgow.

Overall, this case study demonstrated that GenAI can be used effectively to develop innovative assessments to engage students and successfully evaluate intended learning outcomes.

References

Hurren, B L, Rutledge, M and Garvin, A B (2006) 'Team testing for individual success', *Phi Delta Kappan*, 87 (6): 443-447.

Ipek, Z H, Gözümlü, A İ C, Papadakis, S and Kallogiannakis, M (2023) 'Educational applications of the ChatGPT AI System: a systematic review research', *Educational Process: International Journal*, 12 (3): 26-55.

Noy, S and Zhang, W (2023) 'Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence', *Science*, 381: 187-192.

Owen, D and Kay, J (2025) *Transforming higher education learning, assessment and engagement in the AI revolution: the how*. Oxford: HEPI. Available at: www.hepi.ac.uk/2025/07/14/transforming-higher-education-learning-assessment-and-engagement-in-the-ai-revolution-the-how/

Putwain, D W (2007) 'Researching academic stress and anxiety in students: some methodological considerations', *British Educational Research Journal*, 33 (2): 207-219.

von der Embse, N, Jester, D, Roy, D and Post, J (2018) 'Test anxiety effects, predictors, and correlates: a 30-year meta-analytic review', *Journal of Affective Disorders*, 227: 483-493.

Williams, A (2025) 'Integrating artificial intelligence into higher education assessment', *Intersection: A Journal at the Intersection of Assessment and Learning*, 6 (1): 128-154.

Transforming the university learning experience and enhancing student success through authentic assessment and feedback

Frantzeska Kolyda, University of Westminster; Johnathan Lucky, Sky; Helena Disu, University of Westminster

Background

As the world becomes increasingly complex, rapidly evolving and interconnected, we need to prepare our students to be ‘fit for the future’ and to be able to navigate through an AI-enabled world.

At the University of Westminster, our Education Strategy 2023-29 core objective is to “offer personalised and authentic education, underpinned by an inclusive curriculum, to enable all our students, from all backgrounds, to engage in transformative learning and to succeed in their studies and professional lives.” Our strategic commitment is “to develop an outstanding teaching and learning experience that is authentic, practical, active, inquiry-focused, and socially engaged. We are offering authentic teaching, learning and assessment modes which immerse students in the wider world through applied projects, work-based learning and global opportunities” (University of Westminster, 2023, 6). For assessment to remain relevant and future-focused, there is an ongoing need to adopt a holistic approach and reimagine the way we design this.

Assessment and feedback shape the learning experience and – if appropriately and successfully designed – foster a growth mindset in students (Lewis et al, 2025). Student beliefs about learning, challenges, failure and success play a crucial role in their ability to navigate the transitions inherent in the university experience (Kolyda, 2023). According to Self-Determination Theory, developed by Ryan and Deci (2000), when three innate psychological needs are met – autonomy, competence and relatedness – people are most motivated. It assumes that people are inherently inclined toward psychological growth and integration, which leads to learning, mastery and connection with others. However, these proactive human tendencies are not automatic; they require supportive conditions to thrive (Ryan and Deci, 2020). People need to feel that they are effective (need for competence), that they have agency (need for autonomy), and that they experience meaningful connections with other people (need for relatedness). As educators, it is important to consider this when designing learning environments and particularly assessment and feedback. In both approaches discussed here, several elements are incorporated into each assessment to address these three needs.

In the modules examined in this case study, for example, students are required to develop and submit a group agreement as part of the group assessment. This agreement aims to support students in their learning and address any potential challenges. Group work contracts are effective in reducing student anxiety and improving group dynamics and communication, resulting in an improved learning experience overall (Brannen et al, 2021). Students feel a sense of autonomy, moving away from a compliance-based to a collaborative, self-governing approach. They assume collective responsibility, foster a shared commitment and develop negotiation skills that cultivate trust. Clear expectations reduce ambiguity and increase perceived capability.

The learning and teaching methods are inclusive, with practical, active and inquiry focused learning activities designed to support students in their engagement with the assessment and the learning process. In this way, the assessment is placed at the very heart of the module and becomes its central component. As a result, working towards the assessment leads to skill development as opposed to focus only on the product (i.e. output).

Approach

This case study examines two approaches to authentic assessment and feedback at Level 5 in the Computing curriculum, both designed and led by the same module leader, and highlights practice-led insights, challenges, student feedback and reflections.

Approach I: authentic assessment and human-centred design

The Human Computer Interaction and User Experience module aims to provide students with an understanding of how we develop products and systems that provide a successful user experience and support the way people communicate and interact in their everyday and working lives. The module focuses on bringing together the key theoretical foundations of human experiences when people interact with technologies, as well as on the practical skills needed for the design and evaluation of interactive products, systems and services using a human-centred approach.

The assessment provides students with an opportunity to develop and integrate various skills, including problem solving, critical thinking, systems thinking, collaboration and communication. The aim is to empower students to use and develop their skills to address the 17 Sustainable Development Goals (SDGs), understand sustainability as a dynamic concept and recognise the interconnected nature of social, environmental and economic factors.

The assessment takes students through the full development lifecycle of an interactive system. They begin by identifying a problem with an existing system and proposing a solution. They then design and evaluate an interactive system using a human-centred approach. This is a holistic approach that emphasises empathy, iteration and collaboration – critical skills that our graduates need to have developed for graduate success, employability and becoming sustainable citizens. The assessment is closely aligned with real-world scenarios and industry practices.

The module focuses on how to design interactive products that are people centred and meet their needs at a given context. Each cohort consisted of approximately 60 students. The assessment involves both group work (weighting 50%) and individual work (weighting 50%), initial presentation of each group's idea, peer-review, portfolio submission and group presentation.

The assessment redesign involved enhancing the module's original assessment by incorporating an in-class peer review session and by integrating the 17 SDGs as an overarching theme. This peer review component, integrated within group work, aimed to develop students' capacity to give, receive and apply constructive feedback effectively. Peer assessment can prepare students for professional life and is used in a psychological safe environment to provide them with learning experiences and additional feedback from different perspectives that they may wish to consider. It encourages reflective judgment and enhances ownership.

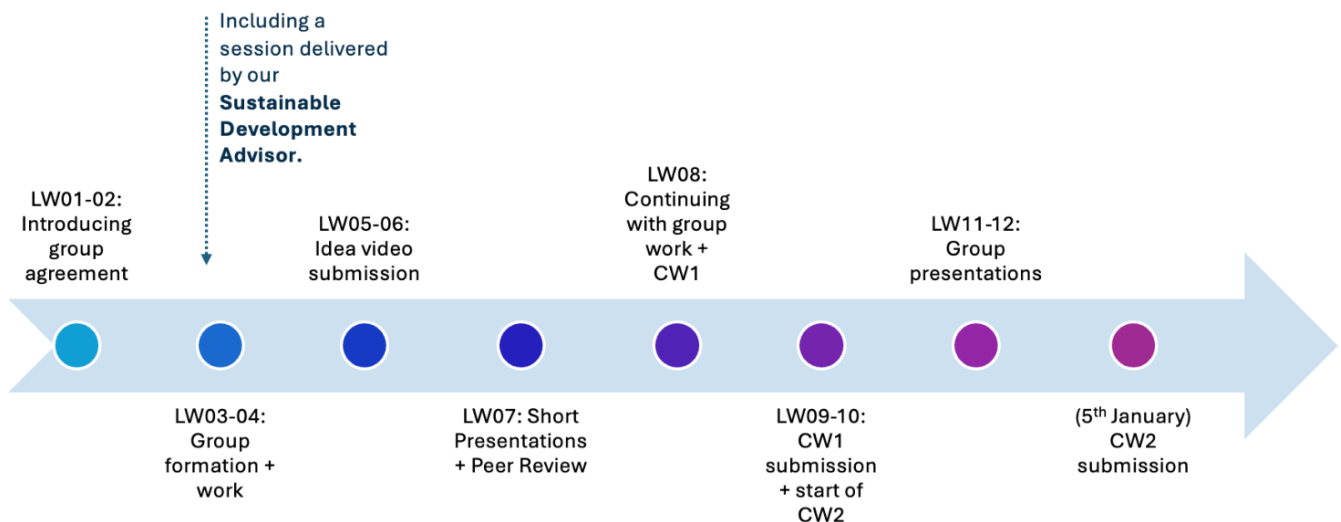


Figure 1. First approach's assessment timeframe

Approach II: authentic assessment and work-based learning

The second approach focuses on a work-based module and a live industry brief. The Agile Project Management and Professional Experience module aims to provide students with an understanding of what agile project management is, respective methods, practices and frameworks and the importance of adapting an agile approach. The assessment consists of an in-class test (weighting 30%) and a group coursework (weighting 70%). Each cohort consisted of 40-50 students.

Our team (comprising the Module Leader, our Industry Exchange Project Officer and our Industry Partner – Delivery Manager at Sky) worked together to create a live group project where students had to develop an agile mindset and skills and propose a solution (i.e. a proof of concept/prototype), to a real business problem that Sky engineering teams were

facing. Students were immersed in a live situation and problem solving and this encouraged them to develop their knowledge and skills to complete the project. The assessment involved groupwork, group presentation and portfolio submission. The students worked in groups of approximately five members and the project was introduced in Week 5. We ran three workshops across Weeks 7–9, led by our Industry Partner, providing students with the opportunity to work on the project. In Week 10, groups presented their work and ideas in class and obtained constructive formative feedback from our Industry Partner and the Module Leader. Students then had two further weeks to refine and complete their project before the formal portfolio submission in Week 12.

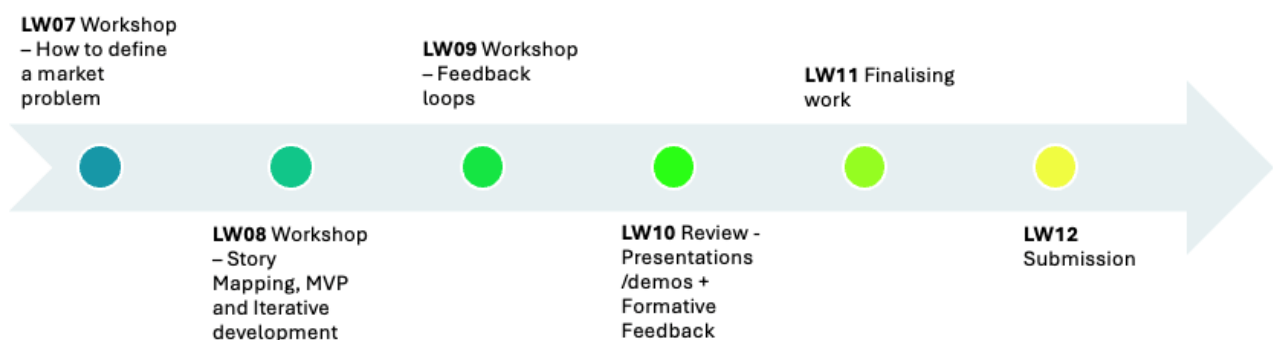


Figure 2. Second approach's assessment timeframe

The setting was a unique opportunity for students to participate in authentic learning and to obtain valuable experience in work-based learning. It was also a great example of how employability could be best embedded in modules and programmes.

The assessment involved redesigning the original group project, placing greater emphasis on the process within a shorter submission timeframe and incorporating direct industry involvement. The redesigned structure centred on three workshops, enabling students to actively engage in collaborative group work, apply their learning in real-time and develop skills directly relevant to industry practice. Previously, longer deadlines often led students to prioritise other module assessments, limiting the potential for sustained engagement and meaningful skill development.

Outcomes

Student feedback for both approaches has been excellent, and their testimonies demonstrated how much students valued authentic learning opportunities and how these made them more determine to learn, develop and grow professionally.

"I like the chance to have group work with other students at our assignments because I think it a good chance to enhance my ability for cooperation. And I also like to finish the assignment without essay but some practical work."

Human Computer Interaction and User Experience Student feedback

Notably, a study abroad student wrote after returning home to explain how the module contributed to securing an internship at a leading technology company. Another student attributed significant academic and personal impact to the module, noting that the learning extended beyond content acquisition to a deeper sense of engagement and meaning. The approach was described as transformative, leaving a lasting impression on the student's overall educational journey.

The experience has been invaluable to our students and fully aligned with our employability provision, helping our students to be 'fit for the future' in the most challenging, cutting-edge labour markets and economic environments:

"I learnt how to be more professional and how to work better as a team. This module is definitely a must for all students at the university. Thank you..... and ... for this opportunity!"

Agile Project Management and Professional Experience Student feedback

"I like the fact that we are doing workshops with Sky and hands-on experience with ... the outside work life environment is like. I find this module interesting to learn and I enjoy learning more about this module."

Agile Project Management and Professional Experience Student feedback

Adopting a holistic approach to the assessment and feedback design ensures that students develop practical skills, engage in authentic learning and integrate key competencies such as problem solving, creative and critical thinking, teamwork and collaboration, as well as presentation and communication.

Both approaches have been designed in ways that enable students to develop a sense of belonging and community as the students have opportunities to work in small groups. Students can develop understanding of diversity while working or interacting with their peers from the same or other courses. Specifically in the first approach, students can develop the acquisition of compassionate capabilities – knowledge, understanding and skills for cultural competence, emphasising trust, empathy and respect. This is possible via practical activities and the assessment, as understanding people (i.e. their abilities and limitations, individual differences, etc) is an important part of the module.

These assessment approaches, grounded in Self-Determination Theory (Ryan and Deci, 2020), include elements transferable to other modules, courses and disciplines, supporting students while creating conditions conducive to learning, mastery and connection with others.

The two strategies followed were examples of how authentic assessment and feedback can be effectively embedded in modules and courses. The case study provides evidence that innovative approaches to authentic assessment and feedback could empower students, helping them develop the skills and confidence to shape their own career paths. Both approaches had a positive impact on the student experience and outcomes. Furthermore,

the study highlights the importance of flexibility, sustainability and authentic engagement. Universities must cultivate safe spaces for pedagogical innovation and the co-creation of bold and meaningful educational experiences (Bryant, 2024), which will be the future focus of this work.

Acknowledgments

The valuable contributions of our Sustainable Development Advisors, Morgan Lirette (2024-25) and Anastasia Lydaki (2023-24), in delivering a workshop on the 17 SDGs to students in the Human Computer Interaction and User Experience module are gratefully acknowledged.

References

Brannen, F S, Beauchamp, D, Cartwright, M N, Liddle, M D, Tishinsky, M J, Newton, G and Monk, M J (2021) 'Effectiveness of Group Work Contracts to Facilitate Collaborative Group Learning and Reduce Anxiety in Traditional Face-to-Face Lecture and Online Distance Education Course Formats', *International Journal for the Scholarship of Teaching and Learning*, 15 (2): 5. Available at: doi.org/10.20429/ijsotl.2021.150205

Bryant, P (2024) *A horizon-scanning report on the changing demographic and pedagogical profiles of current and emerging university students and the responses of the global higher education to the generational challenges*. National Digital Leadership Network Report Series. Available at: doi.org/10.5281/zenodo.15359337

Kolyda, F (2023) 'Fostering a growth mindset in higher education for inclusive learning for all', *Journal of Learning Development in Higher Education*, 27. Available at: doi.org/10.47408/jldhe.vi27.929

Kolyda, F, Lucky, J and Disu, H (2026) 'Enhancing student success, connecting the dots and creating the conditions for all our students to flourish in this ever changing and complex world of work: a case study', in Norton, S and Firth, M (eds) *The 3Es: as easy as A, B, C?* York: Advance HE, pp 27-35. Available at: advance-he.ac.uk/knowledge-hub/3-es-easy-b-c

Lewis, H M, Schneider, K and Tait, D L (2025) 'Using Assessments to Promote Growth Mindset in College Algebra', *College Teaching*, 73 (4): 222-228. Available at: doi.org/10.1080/87567555.2023.2257832

Ryan, R M and Deci, E L (2020) 'Intrinsic and extrinsic motivation from a self-determination theory perspective: definitions, theory, practices, and future directions', *Contemporary Educational Psychology*, 61: 101860. Available at: doi.org/10.1016/j.cedpsych.2020.101860

Ryan, R M and Deci, E L (2000) 'Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being', *American Psychologist*, 55 (1): 68-78. Available at: doi.org/10.1037/0003-066X.55.1.68

University of Westminster (2023) *Education Strategy 2023-2029*. Available at: westminster.ac.uk/about-us/our-university/our-teaching-and-learning/education-strategy

Making research reasoning visible: redesigning a final-year literature review for a GenAI-enabled world

Denise Hough, Lecturer; Victoria Paterson, Senior Lecturer and Barbara K Mable, Professor, School of Biodiversity, One Health and Veterinary Medicine, University of Glasgow, UK

Background

The rapid availability of generative artificial intelligence (GenAI) tools in higher education has intensified scrutiny of written assessment design, particularly where assignments are intended to evidence independent academic judgement. In the life sciences, literature reviews are a core academic practice supporting students' transition to independent research through the development of information literacy, critical synthesis and academic writing. However, conventional essay-based formats tend to privilege the final written product, offering limited insight into how students identify, evaluate and integrate sources, or how they make decisions aligned with research integrity (Nicol, 2010; Winstone et al, 2017). As a result, feedback and support often focus on surface features of writing rather than the research reasoning that underpins effective scholarly practice.

These limitations have become more visible in an AI-enabled context. GenAI tools can generate fluent academic prose, potentially masking students' epistemic agency and amplifying tensions between learning and integrity if assessment responses focus primarily on detection or restriction (Perkins et al, 2023). Sector literature increasingly cautions against such approaches, advocating instead for assessment designs that foreground transparency, judgement and process, and that better reflect contemporary research and professional practices (Sullivan et al, 2023). From a cognitive perspective, this shift is significant: student support must be context specific, discourage passive dependence on digital tools, and encourage active engagement and comprehension, as these capacities are central to critical thinking and shaped by how digital technologies are integrated into learning environments (Kosmyna et al, 2025; Suriano et al, 2025; Rossi et al, 2026).

This case study describes an educator-led redesign of a final-year Life Sciences literature review assessment at a UK research-intensive university. Framed as curriculum enhancement rather than empirical research, the redesign sought to make students' research reasoning more visible by embedding assessment components that surface literature searching, evaluative judgement and reflective decision-making. GenAI was positioned as an optional resource rather than a required skill, giving students agency to decide whether and how to engage with such tools, while receiving explicit guidance to support informed and responsible use. The redesign aimed to normalise transparent discussion of digital tool use, reduce stigma associated with declared GenAI use, and better

align assessment with contemporary research and workplace contexts. The approach draws on principles of authentic assessment (Villarroel et al, 2018) and assessment-for-learning (Nicol and Macfarlane-Dick, 2006), using structured reflection selectively to support metacognitive awareness while avoiding performative or overly procedural accounts (Moniz et al, 2015; Serra et al, 2023).

Approach

Context

The redesign was undertaken within a compulsory final-year Life Sciences course at the University of Glasgow, Animal Biology Core Skills (SCQF Level 10), preparing students for independent Honours research. The course enrolls 100 students annually across Zoology and Marine and Freshwater Biology and has a focus on academic literacy alongside applied statistical skills. The redesign replaced a traditional single-component 2,000-2,500-word essay contributing 40% of the course grade. Students selected one of ten staff-nominated topics that change annually. The redesign was implemented in 2024-25 and refined iteratively in 2025-26 through routine curriculum enhancement.

Design principles

Three principles guided the redesign:

- + the assessment should make students' research reasoning visible to support learning and feedback
- + GenAI guidance should be explicit, contextualised and embedded within assessment design rather than treated as an add-on
- + additional components should support formative dialogue while remaining manageable for markers.

These principles align with research highlighting assessment design as a key determinant of learning behaviour and integrity-related decision-making (Boud and Falchikov, 2006; Bearman et al, 2022).

Redesigned assessment structure

Redesigned materials were collated into an open educational resource, the Assessment Redesign Pack (Hough, Mable and Paterson, 2025), which includes assessment instructions, templates, marking criteria, GenAI guidance and supporting resources. The original essay was restructured into three interrelated components: literature review, annotated bibliography and structured reflective table.

The essay retained its central role in assessing synthesis and critical evaluation but now contributed 80% of the assessment grade. Guidance expanded to include an academic integrity statement and task-specific guidance on GenAI use across anticipated stages of essay preparation, as outlined in our college-level advisory available publicly as an open

educational resource (Hough and Mable, 2025). The assessment brief specified that meaningful GenAI use should be credited in a transparent, non-assessed acknowledgement section, and further clarified that GenAI tools should neither be cited as references nor listed as co-authors, in line with disciplinary standards. It distinguished between grounding evidence-informed scientific writing in primary literature and using broader resources for lower-risk activities such as brainstorming, initial scoping, or developing understanding. Essays were double-marked after briefing markers, with feedback collated into a single form.

The annotated bibliography was introduced to surface students' search strategies and evaluative judgement, enabling feedback on information literacy practices typically hidden in conventional essays (Nicol, 2010). Students annotated the top 20 results from a single database search, indicating if it was included in the final essay with brief justifications. Web of Science was used due to institutional licensing and ease of exporting results into a standardised Excel format. Students were reassured this represented a snapshot of their searching, with multiple searches expected. The annotated bibliography was not graded but penalised if incomplete or not submitted.

The reflective table supported structured reflection across key stages of research and writing. It scaffolded students to describe and evaluate their practices, distinguish descriptive from evaluative reflection, and articulate use of resources and digital tools. This format was selected in response to concerns about the variable pedagogic value of unstructured reflective writing in the sciences (Moniz et al, 2015; Fox, 2025), and to align reflection with stage-specific GenAI guidance. In year one, the marking rubric included a 'scientific rigour' criterion that was removed in year two to reduce performative reflection and focus solely on reflective quality and depth. The reflective table contributed 20% of the overall assessment grade.

Support and implementation

The redesign was supported through scheduled sessions, including an introductory briefing and weekly teaching on literature searching, source evaluation, academic writing and research integrity. Ongoing refinement was informed through routine curriculum enhancement processes, including course feedback, informal staff and student discussions, and anonymous questions raised during teaching sessions. No additional data was collected for research purposes.

Table 2. Reflective Table Template Extract (Timeline column omitted). The body text prompts students for ideas but would be replaced by students' own words.

Activity	Description of your actions (who/what/when & how)	Resources	Reflection on effective process	Okay to use GenAI tools?
Brainstorming, support & collaboration	Why and how did you interact with other people or resources to develop your ideas, approaches or capacity to take on the essay task?	e.g. list resources, individuals, groups, tools	How easy was it to identify and locate the support you needed?	Yes
Initial Reading	How did you build your general understanding of the topic and explored its scope?	e.g. online search, GenAI books, reviews	Any complex topics encountered or identifying ways to refine the scope?	Yes
Databases & Search Criteria	Describe your approaches to structured scientific queries and ensure scientific rigour, e.g. databases, keywords or specific criteria used.	e.g. databases like PubMed, Web of Science, or others	How confident were you in your approach? What sort of adjustments did you have to make and why?	Yes, but not on own to ensure scientific rigour
Selecting Sources	How did you select which sources were relevant, and to what extent did you engage with them?	e.g. any filtering tools or guides	Difficulties in finding, reading, selecting or access to relevant sources?	Yes, but not on own to ensure scientific rigour
Managing Citations & References	Describe how you used tools or implemented methods to organise and keep a record of your references, including correct formatting of the reference list and in-text citations.	e.g. manual or assisted like EndNote, Zotero, Mendeley	Challenges presented by these approaches and how it can improve in future?	Yes, but requires critical review
Drafting	What was your process to put together information in logical arguments for a first draft? How did you keep track of what content came from where and credit it accurately?	e.g. list software, strategies, tools, people	Which aspects were easier or more difficult? Reflecting on your past experiences, is there a pattern in how you approach paper drafting?	No - essential you do this yourself to meet learning outcomes

Assessment and feedback case study compendium 2026

Creating authentic assessments for an AI-enabled world

Proofreading and editing	Outline your proofreading process, which include aspects beyond content creation, such as grammar, spelling or improving flow of information and overall structure.	e.g. software, strategies, tools or Grammarly, people	Any challenges in improving your writing, structuring arguments, etc.? Could you easily identify and locate assistance, if needed?	Yes, but requires critical review
Identify gaps in knowledge	How were you able to get a sense of the gap in knowledge and identify areas where further work is needed to advance this field of study?	e.g. list key resources, people, events	Challenges in narrowing down the gaps in knowledge and think about future directions for studies.	No - essential you do this yourself to meet learning outcomes
Finalisation & Submission Preparation	Describe how you performed final checks, formatting, and submission process, particularly any time management strategies you used.	e.g. list scheduling tools/resources	Challenges in making progress, finalising documents for submission, etc.	Yes

Outcomes

Outcomes are considered in relation to three interrelated areas: visibility of student learning and research reasoning, clarity around GenAI use and academic integrity, and pedagogic insight for staff and future curriculum development. These outcomes draw on routine curriculum enhancement activities rather than formal empirical evaluation.

Visibility of research reasoning and learning processes

A key outcome was improved insight into how students approached literature review tasks, which resources they used, and where difficulties arose. This visibility supported more targeted feedback and helped markers remain informed about evolving research practices, particularly in a rapidly changing digital landscape.

The annotated bibliography revealed patterns in search effectiveness, such as whether many, few, or none of the top search results were used in the final essay, creating concrete opportunities for feedback on search design and source selection. When considered alongside reflective tables, these patterns enabled more individualised support, particularly around database choice, search terms and inclusion criteria.

The reflective tables surfaced conceptual misunderstandings about the literature review process, including conflation of distinct research stages and linear representations of inherently iterative practices. Variation was also evident in students' use of tools and resources, with under-utilisation of reference management software common. Together, the process-focused components provided valuable insight into how students conceptualised and enacted the literature review task.

Iterative enhancement for grading and feedback

Student feedback indicated that articulating research decisions supported greater awareness of disciplinary expectations and increased confidence in navigating uncertainty in academic writing. By shifting attention from written output alone to the reasoning underpinning that output, the redesigned assessment supported preparation for independent Honours-level research.

Some variation was observed across the marker cohort in the depth of engagement with the process-focused components and the feedback provided. However, refinement of the reflective table rubric between iterations improved marking consistency and alignment with the intended learning outcomes. Removing the 'scientific rigour' criterion reduced variability and supported more authentic reflection focused on learning rather than performance. The additional components increased marking time by approximately 10 minutes per student, which staff considered proportionate to the pedagogic benefits.

GenAI guidance, academic integrity and feedback

Embedding task-specific GenAI guidance within assessment instructions, marking criteria and support sessions contributed to greater clarity around expectations for academic integrity. Rather than framing GenAI use as a compliance issue, the redesign foregrounded disciplinary norms, verification of claims and academic accountability. Staff observations suggested that students were more confident discussing their use of digital tools and more able to articulate appropriate boundaries.

Student feedback reflected this clarity, including the quote below. The inclusion of an acknowledgements section further supported dialogue around attribution and contribution, clarifying distinctions between citation, acknowledgement and authorship for both students and staff.

“It was the first time in my degree getting specific advice on the appropriate use of GenAI in scientific writing, finding it refreshing and valuable to have information clearly laid out.”

Final year student (shared with consent)

Implications and conclusion

For teaching staff, access to students’ research reasoning enabled more focused feedback on core learning goals, with the additional insight gained offsetting the modest increase in marking time. Iterative refinements improved consistency and manageability. For students, transparency in expectations not only in terms of responsible GenAI use, but also research practices, was perceived as a benefit of the reflective tasks.

Although situated within a final-year Life Sciences context, the principles underpinning the redesign are transferable across disciplines and levels of study. Assessment designs that make learning processes visible and embed explicit guidance on digital tool use are relevant wherever students are expected to exercise judgement, synthesis and critical evaluation.

Overall, this case study demonstrates how process-focused assessment can respond constructively to GenAI-related challenges without narrowing assessment to detection or restriction. By foregrounding research reasoning, transparency and learning over performance, the approach aligns assessment with disciplinary practice and contemporary research environments, offering a sustainable and adaptable model for assessment in an AI-enabled context.

References

- Bearman, M, Dawson, P, Ajjawi, R, Tai, J and Boud, D (2022) ‘Reconsidering assessment for learning in the age of artificial intelligence’, *Teaching in Higher Education*, 27 (7): 1-15.
- Boud, D and Falchikov, N (2006) ‘Aligning assessment with long-term learning’, *Assessment and Evaluation in Higher Education*, 31 (4): 399-413.
-

- Fox, J (2025) 'Integrating reflection and annotation into writing tasks for science undergraduates', *International Journal of Innovation in Science and Mathematics Education*, 33 (3).
- Hough, D and Mable, B K (2025) *MVLS GAINEd Advisory on UofG GenAI Assessment Scenarios (v1.0)*. Available at: doi.org/10.5281/zenodo.17312534 (accessed 6 February 2026).
- Hough, D, Mable, B K and Paterson, V (2025) *Assessment redesign package: process-focused redesign of an Honours-level literature review in a GenAI-enabled world (v1.0)*. Available at: doi.org/10.5281/zenodo.17456344 (accessed 6 February 2026).
- Kosmyna, N, Hauptmann, E, Yuan, Y-T, Situ, J, Liao, X-H, Beresnitzky, A V, Braunstein, I and Maes, P (2025) 'Your brain on ChatGPT: accumulation of cognitive debt when using an AI assistant for essay writing tasks'. arXiv. Available at: doi.org/10.48550/arXiv.2506.08872
- Moniz, T, Arntfield, S, Miller, K, Watling, C and Regehr, G (2015) 'Considerations in the use of reflective writing for student assessment', *Medical Education*, 49 (9): 901-908.
- Nicol, D (2010) 'From monologue to dialogue: improving written feedback processes in mass higher education', *Assessment and Evaluation in Higher Education*, 35 (5): 501-517.
- Nicol, D and Macfarlane-Dick, D (2006) 'Formative assessment and self-regulated learning', *Studies in Higher Education*, 31 (2): 199-218.
- Perkins, M, Roe, J, Postma, J and Jones, A (2023) 'Artificial intelligence and the future of academic integrity', *International Journal for Educational Integrity*, 19 (1).
- Rossi, S, Fraccaro, V and Manzotti, R (2026) 'The brain side of human–AI interactions in the long term', *npj Artificial Intelligence*, 2: 15.
- Serra, T, Gras, M E, Cañabate, D and Colomer, J (2023) 'Fostering cognitive control through reflection in scientific writing', *Reflective Practice*, 24 (4): 433-446.
- Sullivan, M, Kelly, A and McLaughlan, R (2023) 'Assessment design in the age of artificial intelligence', *Assessment and Evaluation in Higher Education*, 48 (8).
- Suriano, R, Plebe, A, Acciai, A and Fabio, R A (2025) 'Student interaction with ChatGPT and critical thinking', *Learning and Instruction*, 95: 102011.
- Villarroel, V, Bloxham, S, Bruna, D, Bruna, C and Herrera-Seda, C (2018) 'Authentic assessment', *Assessment and Evaluation in Higher Education*, 43 (5): 840-854.
- Winstone, N E, Nash, RA, Parker, M and Rowntree, J (2017) 'Supporting learners' agentic engagement with feedback', *Educational Psychologist*, 52 (1): 17-37.
-

Contact us

All enquiries

Email: enquiries@advance-he.ac.uk

Advance HE helps HE institutions be the best they can be, by unlocking the potential of their people.

We are a member-led, sector-owned charity that works with institutions and higher education across the world to improve higher education for staff, students and society. We are experts in higher education with a particular focus on enhancing teaching and learning, effective governance, leadership development and tackling inequalities through our equality, diversity and inclusion (EDI) work.

Our strategic goals to enhance confidence and trust in HE, address inequalities, promote inclusion and advance education to meet the evolving needs of students and society, support the work of our members and the HE sector.

We deliver our support through professional development programmes and events, Fellowships, awards, student surveys and research, providing strategic change and consultancy services and through membership (including accreditation of teaching and learning, equality charters, knowledge and resources).

Advance HE is a company limited by guarantee registered in England and Wales no. 04931031. Company limited by guarantee registered in Ireland no. 703150. Registered as a charity in England and Wales no. 1101607. Registered as a charity in Scotland no. SC043946. Registered Office: Advance HE, Innovation Way, York Science Park, Heslington, York, YO10 5BR, United Kingdom.

© 2026 Advance HE. All rights reserved.

The views expressed in this publication are those of the author and not necessarily those of Advance HE. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any storage and retrieval system without the written permission of the copyright owner. Such permission will normally be granted for non-commercial, educational purposes provided that due acknowledgement is given. The Advance HE logo should not be used without our permission.

To request copies of this report in large print or in a different format, please contact Advance HE: enquiries@advance-he.ac.uk

advance-he.ac.uk

in **X** **f** **@AdvanceHE**