

“AdvanceHE

+ Enhancing assessment and feedback: a case study compendium

Edited by Stuart Norton, Head Consultant and
Vic Stephenson, Senior Consultant, Advance HE



Contents

Foreword	4
Student perspectives	5
Collaborating with student partners with learning disabilities to close the feedback loop through standardisation	6
Enhancing feedback and assessment consistency in a marking legal essay	14
Sparking curiosity: a student perspective into a statistical poster assessment	20
Enhancing assessment and feedback practices	25
An integrated approach to inclusive assessment	26
A programme-level approach to enhancing assessment practices	32
The how and WHEN of a successful feedback strategy	40
Enhancing assessment and feedback in Master's projects through reflective logs and focus on project skills	48
Supporting students with resit assessments	55
Enhancing student feedback: a semi-automated approach using Microsoft Excel	61
Creating authentic assessments for an AI-enabled world	69
Incorporating artificial intelligence ethically in authentic assessment design and assessment in higher education	70
Authentic assessment design in science – a case study of three approaches	76
Authentic assessment: the current and future role of AI in marketing strategies	81
Maintaining academic integrity in assessment and feedback	87
A multi-pronged and gamified approach to plagiarism prevention and academic integrity training	88
Online maths for social science assessment design in the age of generative AI	96

Enhancing assessment and feedback: a case study compendium

Edited by Stuart Norton and Vic Stephenson

Foreword

Vic Stephenson, Senior Consultant, Educational Excellence, Advance HE

Welcome to our second annual assessment and feedback case studies compendium. The following submissions have been selected from Advance HE's Assessment and Feedback Symposium, held in November 2024, and showcase innovative assessment and feedback practice that makes a tangible, positive impact on the student assessment experience.

Improving student experience of assessment through devising assessments which prepare students for the future, embedding optionality and flexibility into assessment, improving assessment literacy in both students and staff, and ensuring students receive clear, meaningful and actionable feedback has been a priority for many higher education providers for some time. Despite significant and complex challenges presented by advances in digital technology and the changing landscape of higher education, the focus on creating positive assessment experiences has paid off. Students gave assessment and feedback a positive rating of 80.9% in the National Student Survey 2025 – a 2% improvement from 2024, and a 3% increase from 2023. Similarly, Advance HE/HEPI's annual [Student Academic Experience Survey](#) 2025 saw some of the highest ratings for assessment and feedback since the survey began in 2005. Respondents highlighted the speed and usefulness of feedback, the willingness of staff to have further discussions about work, and the quality of the feedback on their progress as particular areas of strength in their programme. This progress is only made through continuous reflection, consultation and collaboration with students and external stakeholders and sharing what works with others.

This collection aims to build on existing scholarship, frameworks, tools, models and examples of best practice and to showcase and discuss initiatives currently being tested across the HE sector, in a wide variety of institutions and disciplines. Although we have presented the case studies under broad themes – student perspectives, enhancing assessment and feedback practice, creating authentic assessments for an AI-enabled world, and maintaining academic integrity in assessment and feedback – it should be noted that many of the approaches have implications across multiple themes.

Each case study is presented in a simple format. They begin with a brief context for the innovation, the contributor then shares their approach, and the measurable outcomes achieved. The case studies include learnings from their own experience, and guidance for others seeking to replicate the approach. We hope the case studies prompt reflection on your own practice and spark discussion on potential enhancements in your assessment practice. We'd love it if you were inspired to share your own innovative approaches at the next Assessment and Feedback Symposium!

Student perspectives

Collaborating with student partners with learning disabilities to close the feedback loop through standardisation

Sharon Sinclair-Graham, Associate Professor in Law, Deputy Head of School and School Director of Teaching and Learning, University of Reading, UK

Stage one: background

The School of Law has long used a standard feedback rubric for evaluating student work across various formats. However, with the shift towards more diverse methods of evaluation beyond traditional coursework and exams, it became clear that the existing rubric was no longer fit for purpose. This is especially true given the increasing emphasis on diverse and inclusive education practices.

An even more pressing concern was the students' response to academic feedback. The feedback loop, an essential part of the educational process, (Biggs 1999, 23) seemed to be fraying. Students were vocal about their dissatisfaction with the feedback provided, citing its often generic nature, lack of actionable insights and sometimes delayed delivery. For students with learning disabilities or neurodivergent conditions, these issues were even more pronounced. Traditional feedback mechanisms were failing to address the nuanced needs of all learners, exposing a critical gap in our approach to fostering an inclusive educational environment.

The project emerged not only from these critical observations but also from qualitative insights gathered through student feedback on module evaluations and NSS scores. Students consistently voiced their dissatisfaction with the clarity, quality and timeliness of feedback. In some cases, they found it to be disconnected from the marking criteria, making it difficult to understand how to improve. For students with learning disabilities or neurodivergent conditions, these challenges were even more pronounced, reinforcing the need for a more structured, inclusive and student-centric feedback system.

Reflecting on these challenges and drawing on Kolb's (1984, 30) experiential learning structure, we recognised that the feedback system did not sufficiently support students' diverse learning needs. This prompted a critical evaluation and overhaul of the feedback process. Our goal was to make feedback more meaningful, timely and actionable for all students, transforming it into an empowering tool for navigating their academic journey. This effort aimed to improve the feedback process and foster a culture of continuous improvement and mutual respect between students and educators.

Students often adopt a passive stance when receiving feedback (Carless and Boud, 2018, 1317) seeking meaning behind their marks yet frequently remaining dissatisfied with the limited depth and utility of the feedback provided. The persistent criticisms revolve around

the quantity and quality of feedback, its timely delivery, and the practical application of the feedback received. The task of the project group was to critically examine how feedback could be restructured to enhance student engagement and how staff could refine key areas to ensure that feedback is both meaningful and actionable. This introspective effort aimed not only to enhance the feedback process itself but also to empower students to engage with it proactively and integrate it effectively into their academic journey.

The University's 15-day assessment return policy is not unusual in higher education. However, this wait can cause student anxiety, leading them to focus on the numeric grade instead of the feedback. As a result, many disregard the feedback, fixating on the mark and potentially submitting complaints without fully understanding the feedback.

Stage two: approach

In 2023, the staff lead secured two Student Partners (each allocated 40 hours of paid work) and received funding for seven additional students through a small grant scheme from the University, amounting to £2,500. This ensured that all students involved in the project were compensated for their efforts. The recruitment call was open to the entire student body, with priority given to those diagnosed with learning disabilities or neurodivergence.

Reflecting on our journey, the project's core purpose was to work alongside students to identify needs and frustrations with feedback. They desired feedback that was timely and insightful, offering actionable steps for growth. The project aimed to implement new rubrics for the 2024-25 academic year.

Three dedicated staff were appointed: the School Director of Teaching and Learning (author and staff lead for the project), the LLB Law Programme Director, and the Staff Disability Officer. The students designed an internal survey to capture candid comments and analyse feedback on students' expectations and experiences with feedback. The project group met five times over the 2023-24 academic year, diligently working on rubrics to cover all assessments; the students consistently steered the discourse. Staff expended approximately 40 hours between them on the project, with the author/staff lead accruing about 25 hours in total.

Survey overview

Nearly half of the students surveyed (n20) found the feedback they received to be inadequate. Their comments highlighted the following issues:

- + 'good' on a sentence does not explain what good is. Tell me what is good
- + feedback is so vague
- + three-word response to a six-page essay
- + 'minor errors but nothing major' – how does that help?
- + 'this needs improvement' – how?
- + not quite there yet.

It was evident that:

- + 50% of students felt feedback rarely provided areas for improvement
- + 60% of students felt there was no clear feedback on structure, legal theories, or case law
- + 50% of students said they would use feedback on their next assessment if it was constructive
- + 55% of students stated that feedback was not constructive, clear, or consistent.

Despite the small sample size, the insights echoed the broader qualitative feedback from our own students in the project group: students felt profoundly let down by the quality and substance of feedback in their assessments. One of the pivotal revelations from this process was the realisation that a one-size-fits-all approach to feedback was inadequate; the School had assumed, perhaps too readily, that the existing rubric was fit for purpose. The diverse nature of assessments, ranging from reflective portfolios to podcasts, necessitated a more tailored feedback rubric. Thus, we envisioned a menu of rubric titles, each specific to different assessment types. While there would be mandatory fields for all staff to complete, the remaining criteria would be tailored to each assessment.

This introspective endeavour was more than just an exercise in improving feedback; it was a transformative experience that underscored the importance of collaboration, inclusivity and continuous improvement. By actively involving students in the redesign process, we not only addressed their specific needs but also fostered a culture of mutual respect and empowerment. This was a poignant reminder that education is a dynamic, evolving partnership between educators and students.

Importance of feed-forward

Students submit their work with the anticipation of receiving constructive feedback they can reflect on – feedback intended to build a scaffold for future assessments. Whether it is formative feedback, guiding students through the learning process, or summative feedback, ensuring they do not repeat the same errors in future assessments, the most crucial component is feed-forward.

Feed-forward is essential as it elucidates how the work should be crafted and informs the student's future endeavours. This element of feedback transcends specific assignments or modules, bearing relevance to all their future academic work. Simply put, the figure below illustrates how students should engage with feedback: they must absorb it, reflect upon it, and apply it in their subsequent assessments. The figure is loosely based on Kolb's (1984) experiential (reflective) learning model, noting that the student can develop through both feedback and lived experience. Kolb's model takes the concrete experience (assessment), moves to reflective observation (feedback), then to abstract conceptualisation (using the feedback), and finally leads to active experimentation (subsequent assessment) – completing a full circle.

This process is not just about immediate academic improvement; it embodies a broader philosophy of continuous growth and adaptation. By internalising and acting on feedback, students develop a deeper understanding of their subject matter and enhance their skills across the board.

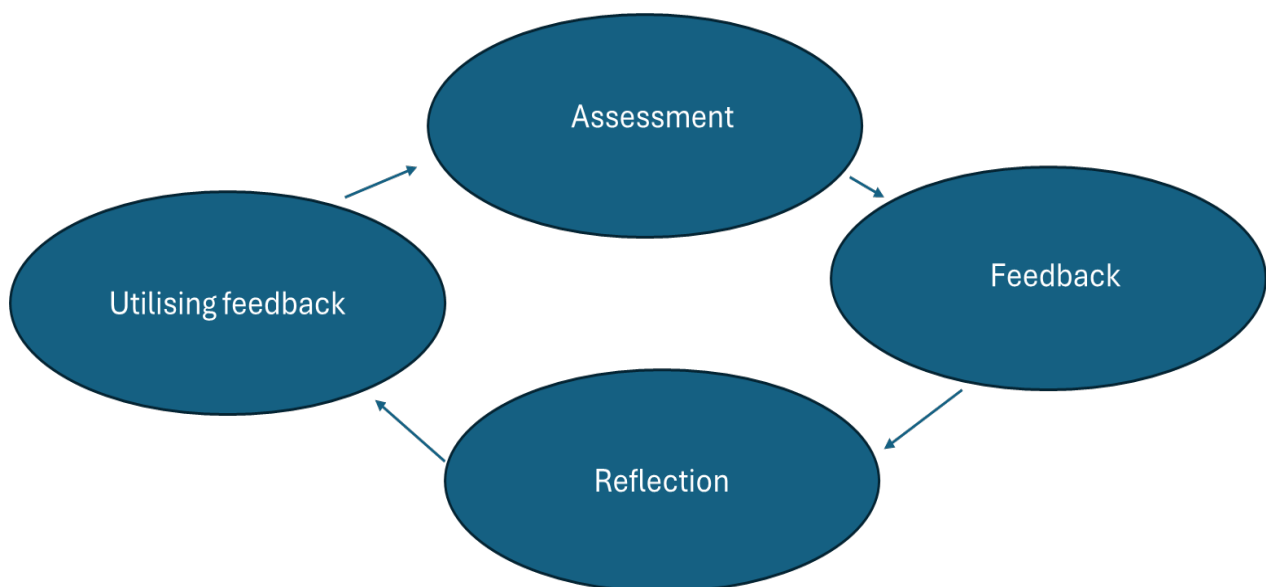


Figure 1. The purpose of feedback

Stage three: outcomes

In the summer of 2024, the culmination of our efforts was realised as the project group compiled a comprehensive set of rubrics tailored to every assessment type within the School. The students, having played a pivotal role throughout, officially endorsed the project before passing it on to the School Director of Teaching and Learning for implementation. These rubrics, which include mandatory boxes (illustrated in Figure 2), require all staff to provide specific feedback when marking assessments. In addition, bespoke boxes were created to address the unique needs of each assessment type. Our Support Centre, responsible for administering the Blackboard pages, then diligently set up these specific rubrics for each assessment.

Status	Rubric Box	What needs to be included in box	
M	Marker and Student Hours - FULL NAME and GRADE	Name, Grade Given and Student Hours	All assessments
M	Moderator - FULL NAME	Confirmation of Grade (ie AGREED) and FULL NAME	All assessments
M	Areas to Commend	What did the student do well - (summary - bullet points)	All assessments
M	Feed Forward - Areas for Improvement	Where can the student improve (summary - bullet points)	All assessments
M	Disagreement with Marking	Please speak to the marker of your assignment first and discuss the assignment with them. If you still do not agree the appeal process is below.	All assessments
		You need to establish a ground for your complaint, and it needs to be one of the following	
		1. Bias by the marker	
		2. Procedural irregularity	
		3. Material defect - either in the teaching or materials.	
		You cannot appeal on the grounds of academic judgment.	

Figure 2. Mandatory boxes

The work undertaken by the project group stands as a model that could be replicated across various Schools within the University. While subject-specific criteria will inevitably vary, the introduction of generic boxes, such as feed-forward and commendation, has proven universally beneficial.

The rubric was implemented in the first semester of the 2024-25 academic year, and student feedback has been overwhelmingly positive. The inclusion of a feed-forward box has been particularly well-received, addressing their primary frustration with previous rubrics which offered little guidance for improvement.

From the staff perspective, the transition has not been without its challenges. As with any new initiative, some staff members have struggled to engage fully, and duplicated boxes in

the rubric were identified. Figure 2 shows the revised mandatory rubric, reflecting these adjustments.

The impact of the rubric will become clear in the NSS results for 2024-25, but we are optimistic that responding to student concerns will yield positive outcomes. In June 2024, a student partner and the School Director of Teaching and Learning presented these developments at a university conference, (later at the Advance HE Conference) sharing our journey and improvements. Their presentation resonated with many attendees, who expressed interest in adopting similar measures in their own departments and schools. The students involved felt empowered by the process, gaining confidence and witnessing the value of their contributions.

STUDENT PARTNERS **University of Reading**

Working in Partnership to Close the Feedback Loop via Feedback Standardisation (School of Law)
Alina Katzighera (Part 3 Law) Josh Robinson (Part 3 Law) and Sharon Sinclair-Graham (Associate Professor in Law)

About the Partnership Project
As Student Partners, we have been working alongside Sharon-Sinclair Graham, other academic staff and TLEP Student Partners for the School of Law since Autumn Term in 2023.

FEEDBACK WORKING GROUP SCHOOL OF LAW

Partnership Project Aims

- Working in a team to analyse feedback from different Law modules to determine what constitutes excellent feedback and the converse
- Reviewing the impact of feedback on the overall student experience with respect to coursework and assessments
- Modelling a feedback template which can be incorporated by academic staff
- Standardising the feedback rubric on Blackboard to produce a harmonised feedback structure

Partnership Project Details

Feedback Working Group Workshop 1

- Introductory meeting between student partners and academic staff for discussions pertaining to delivery of focus groups
- Discussion on the preliminary information and main aims of the feedback working group project
- Identified our collective approach on conducting the overall projects and subsequent workshops
- Reviewing student surveys to obtain general student feedback
- Emphasising disability and neurodiversity considerations to be incorporated in the proposed feedback model
- Require both Student Partner teams to conduct research on feedback as a general pedagogical good practice

Feedback Working Group Workshop 2

- Collated research findings from both Student Partner teams
- Review and discussion of research findings and analysis of additional suggestions on the standardised feedback rubric
- Discussion on conducting focus groups on Week 4 (Spring Term) and chosen students to provide input

Feedback Working Group Workshop 3

- Final discussions on the proposed feedback rubric to be implemented in Blackboard
- Reviewed feedback received from students in their coursework and assessments
- Suggested changes to feedback model for improvement and compiled the information into a Word and Excel document

Reflections

- Undertaking this partnership project has improved our communication skills and promoted a more open dialogue between us (Student Partners) and the academic staff
- We were able to work collaboratively and efficiently in this project with the massive amount of support and guidance from the academic staff and TLEP Student Partners
- This project boosted our confidence to take feasible initiatives to provide an enhanced academic experience to students via our input towards the standardisation of feedback

Figure 3. Student-staff presentation

Key findings and the future

Reflecting on student feedback, incorporating feed-forward mechanisms has significantly impacted the students' academic journey. The call for constructive, forward-looking feedback emphasised not only enhancing their academic skills but also their overall development as learners. Feed-forward fosters a deeper understanding of strengths and areas for improvement, promoting a meaningful connection between students and staff. As demonstrated by Saeed and Mohamedali (2022), this approach nurtures a collaborative and supportive educational environment benefiting both parties.

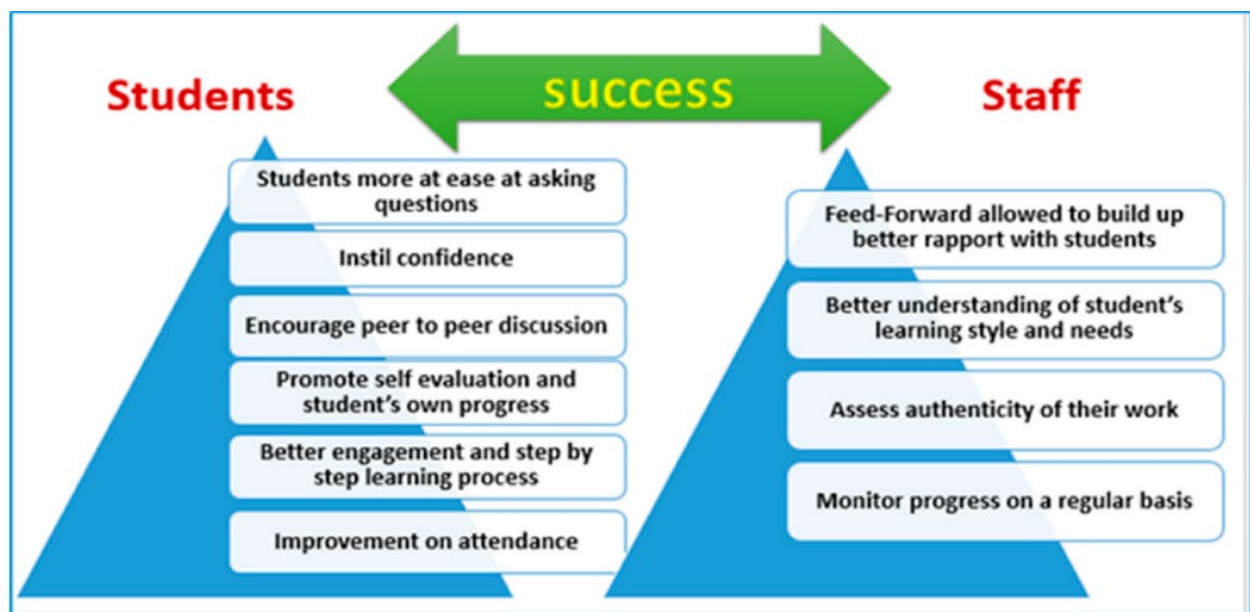


Figure 4. Saeed and Mohamedali – feed-forward impact on students and staff

Ultimately, the project's success stands as a testament to the power of reflective practice and the willingness to embrace change. It highlighted the need to listen, adapt and innovate to foster a more inclusive and supportive learning environment. As we move forward, the lessons learned from this project will continue to guide, inform and inspire our ongoing efforts to enhance the educational experience for all students.

References

Biggs, J B (1999) *What the Student Does: Teaching for Quality Learning at University*. Buckingham: Open University Press.

Boyle, L (2018) 'What has TESTA ever done for us?' University of Dundee Presentation. Available at: <https://www.qaa.ac.uk/docs/qaas/focus-on/what-has-testa-ever-done-for-us.pptx> [accessed 21 February 2025]

Carless, D (2006) *How Assessment Supports Learning: Learning-oriented Assessment in Action*. Hong Kong: Hong Kong University Press.

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. <https://doi.org/10.1080/02602938.2018.1463354>

University of Dundee (no date) *TESTA: Quality Framework*. Available at: <https://www.dundee.ac.uk/qf/learning-and-teaching/testa/> [accessed 21 February 2025].

Hewett, J (2006) 'Formative feedback to students: do we know what we're doing?', in *Fifth Annual International Conference on the Scholarship of Teaching and Learning (SoTL) 2005*, 12th-13th May 2005, Goodenough College, London, p 269.

Kolb, D A (1984) *Experiential Learning: experience as the source of learning and development*. Englewood Cliffs: Prentice Hall.

Saeed, N and Mohamedali, F (2022) 'A Study to Evaluate Students' Performance, Engagement, and Progression in Higher Education Based on Feedforward Teaching Approach', *Education Sciences*, 12 (1): 56. <https://doi.org/10.3390/educsci12010056>

Enhancing feedback and assessment consistency in a marking legal essay

Sau Wai Law, Lecturer in Law, University of Reading, UK

This case study explores challenges in providing consistent grades among different markers and delivering consistent feedback and grading in an assessment under a Year 3 Law module, where students draft briefing notes to simulate legal practice. Issues include generic feedback, a lack of clear vision on how guidelines are implemented, and the difficulty of standardising open-ended assessments. Efforts such as second marking and thematic feedback improved specificity but failed to resolve grading inconsistencies. Long-term solutions may include refining criteria, student involvement and technology-assisted marking. However, these can only partially address the inherent imbalance between flexibility and a consolidated structure, which is unlikely to be governed systematically.

Stage one: background

A Year 3 Law module of an LLB degree programme includes a writing task where students must produce a briefing note. This is a form of authentic assessment that closely mirrors the tasks students will encounter in actual legal practice, where junior lawyers succinctly summarise complex legal issues and provide advice for senior lawyers who rely on these notes to make decisions and advise clients. Students are given five weeks to read a bundle of materials and produce a five-page note.

However, despite the practical relevance of this assessment, significant issues have arisen regarding the consistency of feedback and marks. Students have expressed dissatisfaction with the feedback they receive, which is often generic and unhelpful. Comments such as "it is unclear" or "more analysis is needed" are frequently encountered. Moreover, students receiving similar comments are awarded different marks, leading to frustration among those who struggle to understand the rationale behind their grades.

During a re-marking process, a significant discrepancy in marks between two markers was observed (see Table 1). This discrepancy raised concerns about the reliability and fairness of the assessment process, as well as the effectiveness of the assessment guidance in ensuring consistent grading.

Previous marker	New marker
68↵	52↵
72↵	72↵
58↵	42↵ ↵
52↵	52↵
45↵	55↵ ↵
45↵	32↵
52↵	48↵ ↵
52↵	45↵
52↵	65↵
65↵	72↵
35↵	35↵
45↵	32↵
75↵	55↵
62↵	45↵
62↵	64↵ ↵
42↵	42↵

Table 1. Mark discrepancy between two markers

Theoretical framework

The challenges identified in this case study reflect broader issues in higher education regarding feedback and assessment consistency. Research indicates that effective feedback should be specific, actionable and timely, allowing students to recognise their strengths and areas for improvement (Nicol and Macfarlane-Dick, 2006). However, in disciplines such as law, where responses are often open-ended, providing consistent feedback across diverse answers – especially when there is no definitive ‘correct’ answer – poses significant challenges. Markers must exercise their judgment in interpreting various responses, which can lead to discrepancies in evaluations based on the content presented. This inherent flexibility complicates the establishment of a unified marking guide or model answer that is transparent, consistent and fair, particularly given the rapidly evolving demands of the profession.

While imposing restrictions on the use of specific case law and legislation could foster a more standardised grading approach, it may simultaneously inhibit students' abilities to

demonstrate critical thinking and creativity. This trade-off underscores the difficulty of reconciling consistency with the open-ended nature of legal problem-solving. For example, a student might be penalised for employing an unconventional yet relevant case to support their argument – despite a thorough analysis – simply because it was not listed among the ‘preferred’ cases provided by the marker.

For example, Weber et al (2025) explore the effectiveness of detailed formative feedback in enhancing legal writing, implicitly highlighting the importance of all forms of assessment feedback, including summative grades. However, the relationship between written feedback and grades is not systematically examined in their studies. This disconnect is critical, as effective feedback is essential for student learning, while marks serve as important indicators of students’ performance and abilities. The purposes of receiving a mark and obtaining feedback are distinct but the interrelationship between the two is still overlooked, and its consequences have been reflected in this case study.

Stage two: approach

Actions taken to address significant discrepancies

The initial step undertaken was to investigate the underlying reasons for the observed inconsistencies in marking. This involved a comprehensive analysis of the marking process to identify potential sources of variation. The process involved inviting an additional marker who spent one week re-marking the 16 involved papers. The marker took 15 minutes to mark each paper without referring to any comments or marks from the original markers. Two primary factors emerged from this examination.

First, the inherent ambiguity present in legal essays often permits a range of interpretations, as these essays typically do not yield a single ‘correct’ answer. This ambiguity can lead to discrepancies in grading, as markers may assess papers differently based on their individual interpretations. Furthermore, a significant challenge arises from the lack of clarity regarding the specific interpretations employed by markers; once markers exercise their discretion, the nature of that discretion remains undocumented, complicating any evaluative efforts.

Second, the absence of stringent formatting requirements grants students considerable freedom in structuring their responses, which can further contribute to inconsistencies in evaluation across different markers.

Despite these insights, the establishment of clear criteria has proven challenging, largely due to the complex nature of legal essays. While guidelines can aid in minimising subjectivity and enhancing transparency in grading, they must be crafted with care to avoid being perceived as model answers. Overly prescriptive guidelines risk inhibiting the creativity and critical thinking that are essential to legal education. For instance, a marking guide that mandates adherence to a rigid structure – such as requiring ‘three case law examples’ in every response – may deter students from exploring alternative arguments or innovative approaches.

The second step involved introducing a second marker to evaluate the papers in question, rather than relying on a moderator. While this approach aided in identifying some

discrepancies, it proved time-consuming and did not fully address the fundamental issues at hand. The module convenor continued to rely heavily on the integrity of both markers, and the process fell short of providing a clear solution to the core problems.

In the third step, a thematic feedback approach was implemented. This method organised feedback around key themes such as 'structure,' 'analysis,' 'writing,' and 'referencing,' supplemented by overall comments. To ensure the feedback was relevant and meaningful, student input was solicited regarding what they wished to learn from their assessments. While this structured approach yielded more specific feedback, it often resulted in considerable repetition. General comments regarding students' strengths and weaknesses frequently overlapped with detailed feedback provided in other sections. Furthermore, not all markers were familiar with the thematic approach, leading to inconsistencies in its implementation. Crucially, this method did not resolve the issue of grading consistency. There remained a significant risk of different markers assigning varying grades for similar feedback provided.

The difference between offering written feedback and placing a mark

In the realm of legal education, understanding student expectations is paramount. Students often seek feedback that is both clear and specific. However, they must also recognise that legal essays do not always yield definitive answers. This duality creates a tension that necessitates careful management. For instance, giving student feedback stating "Your analysis of the case law is unclear" does not explain how it impacts its mark. While it is safe to justify a lower grade, it does not encourage constructive feedback, as students might misinterpret such comments as indicators for achieving higher marks.

Traditional assessment rubrics may not adequately address these issues. Often, these rubrics lack the specificity required for legal essays, using broad descriptors such as 'outstanding,' 'good,' or 'effective,' which can be ambiguous. While this ambiguity reflects the real-world nature of legal assessments, it can lead to inconsistencies in grading. For example, a rubric might instruct students to 'demonstrate strong legal reasoning,' yet fail to define what constitutes 'strong' reasoning, resulting in varying interpretations among different markers.

Lastly, although the thematic feedback approach aimed to provide structure, it frequently resulted in repetitive comments. General observations about student performance often overlapped with specific feedback across various themes, thereby diluting the effectiveness of the feedback and rendering it less actionable. For instance, a student might receive the comment, 'Your structure needs improvement,' in both the general feedback section and the dedicated 'structure' theme, without any additional guidance on how to address the issue. This repetition can hinder students' ability to identify and implement the necessary improvements.

Stage three: outcomes

- 1 Refine marking criteria: the next step involves developing detailed guidelines for each thematic area, while maintaining flexibility for markers. This would include creating nuanced rubrics with clear descriptors for different performance levels. This refinement aims to reduce ambiguity and enhance grading consistency, but it necessitates careful communication for markers to ensure effective application.
- 2 Student involvement in the feedback process: engaging students in the feedback process could foster a better understanding of the criteria and their application. This could include peer review exercises or self-assessment activities. While this involvement can lead to more meaningful engagement, it requires careful facilitation to prevent overwhelming students.
- 3 Investigating technology solutions, such as automated marking tools, may facilitate a more systematic analysis prior to assigning grades. For example, natural language processing (aka NLP) tools could provide initial feedback on common issues, enabling markers to concentrate on more substantive evaluations. However, it is essential to ensure that technology complements rather than replaces human judgment. Additionally, it is important to note that for artificial intelligence to yield meaningful insights, all papers must be marked in order to generate sufficient data to assess the consistency between feedback and grading. This presents a significant risk, as moderators may need to revisit hundreds of papers at that stage, potentially leading to inefficiencies and increased workload.
- 4 Implementing regular evaluation of feedback and assessment processes is vital, particularly through calibration in marking – a practice ensuring consistency by aligning evaluators' understanding of grading criteria. Collaborative workshops, where markers review sample assessments, can reduce subjectivity in disciplines such as law and other humanities, social sciences and interdisciplinary fields that rely on a certain level of judgment. Structured feedback from students and markers further refines clarity, fostering cyclical improvement. Resource-intensive calibration frameworks supported by technology (for example, digital rubrics) can standardise grading across subjects. However, over-reliance on tools risks undermining contextual expertise, especially in nuanced fields such as legal education. Balancing technological efficiency with human insight ensures calibration remains a dynamic, adaptable practice, enhancing assessment equity and resilience across disciplines.

Concluding remarks: the art or science?

Aligning feedback and marking practices is essential for fostering a transparent and equitable assessment system within higher education. Effective feedback not only informs students about their performance but also guides their learning journey, while marking serves as a critical metric of identifying outstanding performance. Achieving excellence in both areas requires a nuanced approach that acknowledges the complexities involved in providing constructive evaluations.

Governance for markers plays a vital role in establishing frameworks that facilitate effective feedback delivery. Clear guidelines enable markers to apply criteria consistently while engaging meaningfully with students. However, monitoring the quality of feedback and marking can be difficult, as this reliance on individual markers can introduce biases and variability in their commitment.

This case study underscores the complexities of providing consistent feedback in subjective assessments, particularly within legal studies. While clear marking criteria can improve uniformity, overly detailed frameworks may inhibit assessors' judgment. Therefore, effective feedback and marking necessitate a balanced approach. By integrating clear criteria with personalised feedback and committing to ongoing evaluation, institutions can create an assessment environment that genuinely supports student learning and promotes academic growth.

References

Nicol, D J and Macfarlane-Dick, D (2006) 'Formative assessment and self-regulated learning: A model and seven principles of good feedback practice', *Studies in Higher Education*, 31 (2): 199-218.

Weber, F, Wambsganss, T and Söllner, M (2025) 'Enhancing legal writing skills: The impact of formative feedback in a hybrid intelligence learning environment', *British Journal of Educational Technology*, 56 (2): 650-677.

Sparking curiosity: a student perspective into a statistical poster assessment

Rachael Ainsley Addy, Senior Lecturer; Ellie O’Grady, Student, BSc (Hons) Animal Behaviour and Welfare, University of Lincoln, UK

Stage one: background

Beginning in academic year 2020, the University of Lincoln introduced a new maths module within its Science Foundation Year (SFY) programme. This was an integrated Year 0 covering multiple science-based programmes across multiple schools. The new module focused on developing, supporting and applying fundamental maths skills for students in three schools (Biology, Chemistry and Pharmacy), as all destination programmes to these schools required maths and statistical skills. Between 2020-24, the module had a peak of around 120 students. Previous maths and study skills-based SFY modules had attempted poster assessments unsuccessfully. However, with a move towards authentic assessment (University of Lincoln, 2022) and a clear link between academic research posters and statistics (Miller, 2007), the module was designed with a group statistical poster summative assessment, using open-source, context-appropriate data.

During the academic years 2020-22, the group statistical poster assessment ran with varied success and logistical difficulties for all parties, with each student submitting their files (meaning sometimes not all members of a group would submit the same files), a maximum of five groups being timetabled at any one time for presentation and a second presentation session organised for groups with extensions. At times this caused confusion over when and where groups should attend and, when presenting, students used small A4 or A3 physical copies of their posters, which were typically ignored during the presentation. Grades were diverse, with many students focused on the research background or poster design over statistical content. Where results and experiences were positive, students reported long-term impact as they progressed into their destination programme. However, where experiences were negative, students reported feeling a lack of worth to the experience. For staff involved, reviewing the diverse quality of work produced and outcomes achieved supported the view of the group statistical poster assessment being an innovative assessment practice (Ashford-Rowe et al, 2014; Swaffield, 2011), but that there were challenges to be overcome if it were to reach full potential. An alternative approach was therefore designed to improve student experiences and skill levels, support better outcomes and develop students with stronger skill profiles when progressing to their destination programme (Brown, 2018). The insights discussed below relate to the 2022-24 academic years.

Stage two: approach

Statistics became the sole focus of Semester B (50% of the module), which was delivered in two blocks prior to the group statistical poster assessment being launched, with the presentation becoming a poster conference requiring A1 or A0 posters. Theoretical knowledge was delivered in weekly face-to-face lectures and asynchronous online materials, before practical and applied knowledge was delivered in computer lab sessions (with the sessions recorded for later review if required). The group statistical poster assessment was streamlined to improve experiences for staff and students, becoming a group submission (one set of files per group) and presentations arranged at least 14 days post submission to allow those with extensions to participate in the same presentation event.

For 2022-23, the group statistical poster assessment remained the only summative assessment for statistics, but in 2023-24 a second summative assessment piece was introduced: individual data analysis. This was adopted to ensure all individuals could effectively apply relevant statistical skills in a context-appropriate way and was supported by student feedback suggesting some groups assigned tasks as statistical/non-statistical, enabling students to circumvent their lack of understanding. While the individual data analysis assessment will not be discussed here, it is important to note that both assessments used open-source, context-appropriate data and the statistical expectations remained constant across both assessments, meaning the statistical skills and processed required did not change, only the context of application and dataset.

Despite the previous challenges with group statistical poster assessment, this continued to be viewed as an appropriate method of assessment for applying and deepening skills and knowledge acquired in both the taught and applied environment. It was also viewed as a positive and flexible option to promote academic and personal development as research posters are authentic (Brown, 2018) to the destinations of students and would allow for use of artificial intelligence. All students had the opportunity to use artificial intelligence if they wished. However, alongside the final poster, each group provided a working file showing their raw dataset, any manipulations to the dataset and all statistical analyses undertaken, which was marked against an expected format. Therefore, while artificial intelligence could not be used as a replacement for any statistical skills required, it could be used for any other aspect, such as design and flow.

The group statistical poster assessment was made known to students from the beginning of the module and discussed in increasing detail leading up to the launch. This three-hour, intensive, face-to-face session with significant content (including peer reviewing, exposure to the rubric and support resources available [both module and non-module-specific]) encouraged deep understanding of the group statistical poster assessment before beginning. Groups collaboratively peer reviewed multiple anonymous submissions from previous cohorts against the current rubric, to consider strengths and limitations, interrogate the nuances of the rubric and evaluate a variety of approaches taken by others. Students were not, however, given a step-by-step guide to follow to complete the group statistical poster assessment, and had free reign of which open-source dataset to choose (from a list

provided), how to use this dataset (including data manipulation) and formulating their own hypotheses. This was difficult for some students, as there were no formal constraints for this and outside of their group it was therefore more difficult to compare with peers, as differing hypothesis and data manipulation would cause varying outcomes. To ensure groups were supported during the process, they had timetabled workshops (providing dedicated time to work as a group, but also access support from an academic should queries or difficulties arise), timetabled drop-in support (providing dedicated time an academic would be available) and weekly open office support (where groups could book one-to-one meetings).

For staff, the changes to the group poster assessment had a positive impact on workload, as marking decreased due to group submissions and the poster conference being held on a date that would include those with extensions, preventing planning and staffing of multiple face-to-face sessions being required. The only additional work required was creating groups within the Blackboard virtual learning environment for group submission and file sharing and ensuring the rubric and open-sources datasets remained relevant and representative. All other supporting information, such as examples of a range of scientific posters, help files, etc, were already in place from the 2020-22 cohorts. In addition, by adjusting the module timetable to place all statistical content into Semester B and increasing the weighting this contributed to the module, students had clear delineation and placed a greater importance on the content. Finally, while application of statistics using computer software had been delivered previously, timetabling face-to-face computer lab sessions allowed students to receive immediate support and feedback, when compared to the previous online synchronous delivery, providing a stronger mechanism for support despite the content remaining the same, improving student engagement and confidence. Overall, once the initial group statistical poster assessment was created, changes reduced demands on staff, while increasing student satisfaction and application of knowledge.

Stage three: outcomes

Notwithstanding the high levels of direct support and supporting information to approach the group statistical poster assessment, there remained individual differences in the approach and comfort levels of students. As expected, there are those for whom statistics was an enjoyable experience and they can visualise the group statistical poster assessment from the outset, but also those who, during the launch session, commented “This is too difficult”, “This doesn’t make any sense” and “I will never use this”, despite having not yet had the opportunity to reflect on the information and formulate a plan with their group. Being aware of these intra-cohort differences and to foreshadow negative commentary, students were repeatedly asked to trust the ‘tips and tricks’ presented by the academic, and that the process does become clearer with engagement, a belief that became reflected in the module feedback. The module feedback, and specifically when considering the group statistical poster assessment, was positively viewed, with student comments including “The poster conference has prepared us for situations in the future”, “The experience we had by doing the poster is the best thing about the module” and “I wish we had more opportunities like this as it was a great experience”.

To better understand the change in student mindset, from hesitancy to fantastic experience, it was important to listen to student voice, one of whom was Ellie. Ellie was studying Animal Behaviour and Welfare, with the view to becoming a zookeeper, and reported challenging previous maths experiences, including initially failing her GCSE maths (passing on resit during Covid) and previously failing this maths module, in her words “Because I just wasn’t very good at it”. Ellie shared her experiences from both efforts at the group statistical poster assessment, explaining that when she initially undertook this the first time, she did not understand the task as she often did not attend, and her group choice enabled her to need not contribute more widely than towards the wider reading, as the topic chosen was one for which she had underlying knowledge. Ellie also explained that her second experience was very different, not only because the group worked together to formulate all aspects, but also due to an understanding of the statistics and the overall purpose. While it could be argued this came from second exposure to content, as already highlighted, by her own admission, Ellie did not attend much of the delivery during her first exposure. Ellie also noted that she had identified the need to seek support each week to ensure she felt secure in her understanding, which had been beneficial across the module in general as this had improved her confidence.

More generally, Ellie reflected that the combination of taught and applied work allowed theory to be put into practice and challenge students to replicate the notes they had made, questioning, “Do you actually understand this, or do you just understand what you have written?” and challenge students with group work, especially when working with friends. Both times Ellie was in a group with friends, but experienced different group dynamics, with one member taking the lead and undertaking the majority of work during her first undertaking, leaving her “...staring at my computer screen for days like I was looking at Spanish and I don’t understand Spanish”, while during her second undertaking there was a feeling of empowerment and confidence from the outset, communication and collaboration was strong and each group member took ownership of a specific part of the overall output. This was reflected in the latter groups’ wider considerations, including a conscious decision to reflect good practice regarding accessibility, which had not been explored with the former group.

Wider feedback from the 2022-24 cohorts (around 344 students) reflected the differences individuals held at the group statistical poster assessment launch regarding what to do with the data and information presented, but that instructions were clear and appropriate to destination programmes once they were able to consider everything outside of the session. Students reflected on the challenges posed by group work to ensure that all members participated in a way that positively challenged their abilities (for example, patience), so while the group statistical poster assessment could potentially be an individual task, there was an understanding of the need to develop collaborative skills for their futures in research-focused destination programmes. Finally, feedback also highlighted the importance of the support sessions (workshops, drop in, open office) for critical thinking and self-confidence with the process, as academics would coach students through difficulties, without providing a solution. This feedback contrasts to the 2020-22 cohorts who had mixed and at times

negative feedback towards the group statistical poster assessment. Due to restructure, it is no longer possible to directly compare proportional changes towards positive feedback across cohorts, nor provide direct quotations outside of those stated.

The most unexpected reported impact of the group statistical poster assessment, however, related to the poster conference itself. Students voiced they felt like academics, and it was a proud moment where they were able to share their work with others (students and academics), answer questions and show how proud and positive they were to have moved from “I didn’t understand this at first” to “Look what has been achieved with hard work”.

References

Ashford-Rowe, K, Herrington, J and Brown, C (2014) 'Establishing the critical elements that determine authentic assessment', *Assessment and Evaluation in Higher Education*, 39 (2): 205-222.

Brown, J A (2018) 'Producing scientific posters, using online scientific resources, improving applied scientific skills in undergraduates', *Journal of Biological Education*, 54 (1): 77-78.

Miller, J E (2007) 'Preparing and Presenting Effective Research Posters', *Health Services Research*, 42 (1p1): 311-328.

Swaffield, S (2011) 'Getting to the heart of authentic Assessment for Learning', *Assessment in Education: Principles, Policy and Practice*, 18 (4): 433-449.

University of Lincoln (2022) *Teaching, Learning and Student Experience Strategic Plan 2022-27*. Lincoln: University of Lincoln.

Enhancing assessment and feedback practices

An integrated approach to inclusive assessment

Eliza Jess Power, Director (Professor) of Learning and Teaching; Davina Whitnall, Academic Development Lead; Maggie Scott, Associate Dean Academic (Quality Assurance and Enhancement), University of Salford, UK

This case study builds on the presentation at Advance HE's Assessment and Feedback Symposium 2024, illustrating how collaborative activities were embedded, drawing on best practices, to develop a practical, evidence-informed set of five principles to enhance assessment and feedback.

Stage one: background

An integrated approach that designs inclusive assessment into higher education (HE) curricula is crucial for achieving the ambitious participation targets set by Universities UK, which aim for “70% of the population aged 25 to be studying at level 4 or above by 2040” (2024, 4-5). Given the current participation and completion rates, this target is challenging and highlights the necessity for institutions to support students in their chosen study paths, regardless of demographic characteristics, including gender and ethnic origin, and irrespective of previous learning experience. It is recognised that sustainable change with increasingly diverse student populations requires evidence-informed best practices to drive effective institutional strategies and inform sector practices (Salford, 2022; Teeside, 2022).

A significant challenge in the sector is the lack of progression for some learners. As Thomas et al (2017) notes, every student who drops out of HE represents a loss to their institution, the future economy and, most importantly, to their own personal development. Therefore, institutions must critically evaluate their practices to understand the journeys of individual students and cohorts and make thoughtful choices to support their needs (Blake et al, 2022). Supporting transition to university and between levels is key to student success, and this case study details a collegial approach to developing institutional inclusive assessment principles.

Fostering an inclusive environment is crucial for ensuring all students have equal opportunities to succeed (Everett and Oswald, 2018; Gawthorpe and Hubbard, 2024). Bishop-Monroe and Garcia (2023) define an inclusive culture as one where all students feel welcome, comfortable and confident, sharing ideas and participating in the educational process. Building belonging requires integration through connectedness and inclusivity (Blake et al, 2022), creating an environment where students are motivated, and can actively participate in their learning process (Gulko et al, 2024; Thomas et al, 2017).

Previous evidence-based research by Power et al (2024) focuses on supporting transitions from vocational qualifications into HE, working with eight academic communities to gather insights into effective practices. These findings reinforce the need for targeted support mechanisms that bolster student confidence (Whitnall et al, 2024). Key findings from this research emphasise the importance of diverse pedagogical approaches, inclusive assessment practices and feedback mechanisms that reflect students' diverse strengths and needs with varying entry qualifications. Our approach to gathering sector best practices in the context of cultural inclusivity to devise an intentionally inclusive approach to assessment and feedback is underpinned by three key theoretical frameworks: Hattie and Timperley's Feedback Model (2007), Nicol and Macfarlane-Dick's Formative Assessment Principles (2006), and Black and Wiliam's Theory of Formative Assessment (2009). By integrating these theoretical frameworks into a staff-student co-created model, the University of Salford has developed a robust approach to inclusive assessment and feedback. The case study illustrates how these principles have been developed from theory into a co-created framework and then practically applied, emphasising the university's commitment to creating an inclusive learning environment that empowers all students to achieve their full potential.

Stage two: approach

Our approach involved an initial review of sectoral practice in the context of cultural inclusivity to then focus on the 'Salford' context to devise an intentionally inclusive lens for assessment and feedback. We brought together a range of institutional best practice and sector examples to explore further. This learning evolved into an internal report and literature review of approaches to develop a foundation for further exploration within the institutional context. Learning from institutional best practice was essential to build an evidence-informed base within our local environment. Care was taken to position the approach to be supportive and not a 'deficit' model. The framework of principles was co-created with students to focus on voice and agency so that lived experiences and a range of perspectives underpinned its development. The three innovations of framework design, collaborative activities and supportive environments not only helped to shape the agenda but fostered a creative and focused space for staff to connect, explore and share.

The workshop approach enabled staff to enhance learning through collaboration and connectivity. In exploring best practices, the utility of pedagogic interventions such as personal tutors, use of language and command words and Universal Design for Learning were shared. Immersive learning spaces across various disciplines fostered a sense of belonging and confidence among learners at Salford.

The insights gathered enabled us to develop a set of principles to guide inclusive and effective assessment and feedback practices. The output was the development of a practical, evidence-informed set of five principles to enhance assessment and feedback. The principles state that each of our assessments should embody the following approaches:

- + clear – we employ straightforward language and processes in our assessment briefs and feedback, ensuring students know what is expected of them throughout their educational journey
- + understood – by simplifying assessment and feedback mechanisms, clearly aligned with module learning outcomes, learning activities and the wider programme, we foster shared comprehension between staff and students, promoting universal assessment literacy
- + authentic – developed in partnership with stakeholders, resonating with students' interests and lived experience, ensuring they are anchored in practical real-world skill application, preparing them for the professional world
- + robust – with opportunities for formative and synoptic assessment of student activity
- + personalised, offering opportunities for diverse activities with flexible assessment options and opportunities for negotiation, allowing students to express their knowledge in various ways, so that assessments cultivate community and ownership.

The goal was to ensure all students benefit from an inclusive assessment strategy. In addition to the principles themselves, key findings and recommendations enabled this work to progress at pace. These included opportunities to:

- + develop support programmes targeted to enhance student confidence while they are transitioning into HE
- + encourage diverse teaching methods to accommodate students with varying entry qualifications
- + foster a culture of inclusive assessment methods and constructive feedback to support all students
- + invest in extra/intra-curricular and community-building initiatives to promote a sense of belonging
- + provide staff training and awareness programmes to value and celebrate diverse entry backgrounds and to assist programme teams in their awareness and understanding of data
- + encourage Level 4 to be staffed with permanent staff to support transition
- + consider the positioning of professional skills modules to provide specific support at all levels of the learner journey and that personal tutoring sessions are formally timetabled.

Stage three: outcomes

We are currently exploring the institutional impact of the principles and success measures. Early analysis indicates:

- + a number of programmes that participated in the original workshops showed improvements in 'Pass all modules first time' data, reducing the gap between students who joined the university from BTEC and A-level entry routes by up to 11% (2021-22 to 2022-23), with further progress being made in 2023-24 in specific programmes up to 10%, with others maintaining the gap closure
- + the gap in achieving good degrees between A-Level and BTEC entry routes narrowed by 4% from 2021-22 to 2023-24
- + the wider impact of this work has yet to be realised, and further data analysis is being undertaken.

We recognise that the project is exploratory and while we gained new insights from this activity which enabled us to formulate robust inclusive assessment principles, we acknowledge that further data interrogation is vital to assess the full impact of this work. However, our evidence-informed approach has opened conversations and demonstrated a shift change in thinking, sharing and contributing to a wider community of practitioners.

Reflecting on next steps, we have begun to build and extend on this work through sharing best practice and initiating further workshops to support programme teams. Furthermore, outputs from the community of practitioners produced a series of artifacts that provide evidence of real impact in the form of podcasts, escape rooms, language toolkits and co-created workshops, feeding into the further refinement of the principles and evidence base. Considering the progress of this work and within the 12-month timescale of the project, a change in practice in moving to an evidence-based approach has begun to transform the institutional approach to assessment. The timeline had moved quickly and supported culturally responsive teaching, beginning with the review of sector practice which began in spring 2023 to the designing of principles and the development and delivery of workshops that took place between May-July 2024. The rapid and agile development process has facilitated a flexible approach that has benefited programmes, creating meaningful action that's responsive to and cocreated by the Salford community.

We are developing further guidance to support our inclusive assessment and feedback principles, drawing on good practice from our academic community. The intention is to build on the recommendations and learning through further collaboration with programme teams to socialise, implement and embed new initiatives to support inclusive assessment. We continue to track progress through the established metrics and are cross-cutting data to assess the wider impact on progression to underpin our university strategy of innovating to enrich lives.

Ethics

Ethical approval was obtained for this study (the University of Salford, Academic Ethics Institutional Review Board. Approval Number:11866). Written informed consent was obtained from all individual participants included in the study.

Acknowledgements

We express our deepest gratitude to all contributors of this study for sharing their rich and diverse experiences. We would like to thank the Levelness Project Working Group and the following individuals: Frankie Baines, Lauren Beckett, Dr Eve Blezard, Philip Boyle, Annette Cooke, Dr Janine Crosbie, John Dermo, Timothy France, Dr Nicola Grayson, Nicole Holt, Dr Dean Irwin, Jayne Langlands, Manisha Mort, Dr Sara Namvar, Dr Christina Patman, Dr Emma Smith, Dr Calum Thomson, Dr Paul Wilson, and Dr Gerard Wood.

References

- Black, P and Wiliam, D (2009) 'Developing the theory of formative assessment', *Educational Assessment, Evaluation and Accountability*, 21 (1): 5-31.
- Blake, S, Capper, G and Jackson, A (2022) *Building Belonging in Higher Education*. Pearson WonkHE. Available at: <https://wonkhe.com/wp-content/wonkhe-uploads/2022/10/Building-Belonging-October-2022.pdf> [accessed December 2024].
- Brown, J A (2018) 'Producing scientific posters, using online scientific resources, improving applied scientific skills in undergraduates', *Journal of Biological Education*, 54 (1): 77-78.
- Everett, S and Oswald, G (2018) 'Engaging and training students in the development of inclusive learning materials for their peers', *Teaching in Higher Education*, 23 (7): 802-817.
- Gawthorpe, P and Hubbard, H (2023) *QAA Collaborative Enhancement Project. The Inclusive Education Framework: Final Evaluation Report March 2023*. Available at: https://www.qaa.ac.uk/docs/qaa/members/the-inclusive-education-framework---evaluation-report.pdf?sfvrsn=3f9aaa81_8 [accessed December 2024].
- Gulko, N, Wood, N, Blondeel, E, Churyk, N T, Derbyshire, L E, Kawor, S, Lento, C, McGuigan, N, Merendino, A, Middelberg, S L, Sahoo, S K, Tong, J T and Withanage, N (2024) *Enhancing Inclusive Student Engagement in Higher Education: Literature Review*. Quality Assurance Agency for Higher Education. 1-22. Available at: https://gala.gre.ac.uk/id/eprint/48661/7/48661%20WITHANAGE_Enhancing_Inclusive_Student_Engagement_In_Higher_Education_%28QAA%20REPORT%29_2024.pdf [accessed December 2024].
- Hattie, J and Timperley, H (2007) 'The power of feedback', *Review of Educational Research*, 77 (1): 81-112.
- Miller, J E (2007) 'Preparing and Presenting Effective Research Posters', *Health Services Research*, 42 (1p1): 311-328.
-

Nicol, D J and Macfarlane-Dick, D (2006) 'Formative assessment and self-regulated learning: A model and seven principles of good feedback practice', *Studies in Higher Education*, 31 (2): 199-218.

Power, E J, Blezard, E, Whitnall, D, Thomson, C, Dermo, J and Grayson, N (2024) 'QAA insight A thematic approach to enhance transition between levels, for students with vocational qualifications', in *Quality Insights Conference*, 22-23 February 2024, Virtual.

Swaffield, S (2011) 'Getting to the heart of authentic Assessment for Learning', *Assessment in Education: Principles, Policy and Practice*, 18 (4): 433-449.

Thomas, L, Hill, M, O'Mahony, J and Yorke, M (2017) *What works? Student retention and success HEA*. Available at: <https://lizthomasassociates.co.uk/downloads/2017/Summary%20report.pdf> [accessed December 2024].

University of Lincoln (2022) *Teaching, Learning and Student Experience Strategic Plan 2022-27*. Lincoln: University of Lincoln.

University of Salford (2022) *The University of Salford's Enabling Student Success institutional case study*. Advance HE Collaborative Development Fund: Inclusive Institutions project. York: Advance HE. Available at: www.advance-he.ac.uk/membership-2021-22/member-benefits/collaborative-development-fund-2022/inclusive-institutions/inclusive-structures-driving-transformational-change [accessed December 2024].

University of Teesside (2022) *Embedding Inclusive Assessment: A Reflective Toolkit*. QAA Collaborative Enhancement Project led by Teesside University. Available at: https://www.qaa.ac.uk/docs/qaa/members/embedding-inclusive-assessment-reflective-toolkit.pdf?sfvrsn=7369a281_8 [accessed December 2024].

Universities UK (2024) *Opportunity, growth and partnership: a blueprint for change*. Available at: <https://www.universitiesuk.ac.uk/sites/default/files/field/downloads/2025-01/opportunity-growth-and-partnership-a-blueprint-for-change-170125.pdf> [accessed December 2024].

Whitnall, D, Scott, M and Power, J (2024) 'An integrated approach to inclusive assessment', in *Advance HE's Assessment and Feedback Symposium 2024*, Virtual.

A programme-level approach to enhancing assessment practices

Grazyna Grace Lipowska-Bhalla, Postgraduate Programme Director, Lecturer in Cancer Biology, University of Manchester, UK

Stage one: background

This case study discusses the programme-wide approach I implemented to evaluate and enhance assessment practices within the Level 7 postgraduate taught programme I lead as Programme Director.

Assessment plays a pivotal role in the teaching and learning process (Biggs and Tang, 2011). It not only shapes students' engagement with their studies but also serves as a key mechanism for maintaining academic standards and ensuring programme quality. As Boud observed, *“Students can, with difficulty, escape from the effects of poor teaching, they cannot (by definition if they want to graduate) escape the effects of poor assessment.”* (Boud, 1998). This underscores the necessity of cultivating an effective assessment environment that supports student learning and enriches their overall educational experience (Gibbs and Dunbar-Goddet, 2009).

The higher education sector has undergone significant transformations in teaching and learning methodologies in recent years. The shift from traditional classroom-based instruction to flexible and blended learning models, coupled with the rapid emergence of AI-driven technologies, necessitates the redesign of assessment practices to ensure they remain relevant, inclusive and authentic. While individual course modules have adapted their assessments to keep pace with these changes, such enhancements have often been implemented in isolation, without coordination at the programme level.

This lack of alignment can result in unintended shifts in assessment patterns, such as the overuse of certain methods, an excessive reliance on summative assessments and the clustering of assessments within short timeframes. Such heavy, clustered and repetitive assessment patterns encourage surface learning rather than deep engagement with course material (Jessop and Tomas, 2017; Tomas and Jessop, 2018; Wu and Jessop, 2018). Misaligned assessments also disrupt constructive alignment, leading to an uneven evaluation of intended learning outcomes (ILOs). Some ILOs may be over-assessed while others are overlooked, ultimately undermining the coherence and effectiveness of the programme's overall learning goals (Biggs, 2003; Bloxham and Boyd, 2007).

A cohesive, programme-level assessment approach is essential in creating effective assessment patterns and shaping student learning behaviours. Well-distributed, diverse and feedback-driven assessments not only foster deeper learning and sustained engagement but also ensure that assessments serve a clear and meaningful purpose within the broader educational framework (Gibbs and Simpson, 2004; Jessop et al, 2014).

Stage two: approach

A structured, programme-wide approach was developed to enhance assessment practices within the Level 7 postgraduate taught programme in Cancer Pathology. This one-year, full-time programme consists of eight 15-credit taught modules across two semesters, followed by a 12-week, 60-credit research project in the summer. To establish a more coordinated and balanced assessment strategy, an Assessment Enhancement Framework (Figure 1) was developed to guide improvements across the programme.

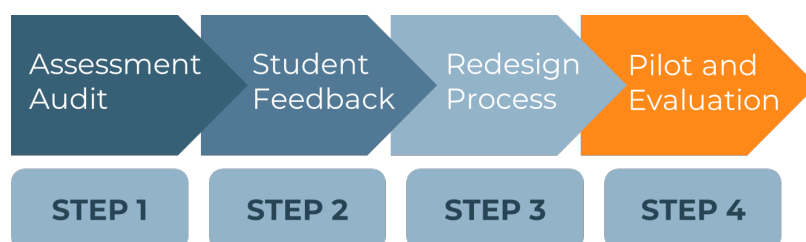


Figure 1. Assessment Enhancement Framework

This case study focuses on the first two stages: evaluating existing assessment practices and collecting student feedback.

A programme-wide assessment audit was conducted during the 2023-24 academic year to evaluate the diversity, volume and distribution of assessments, as well as their alignment with programme ILOs. The audit was informed by a comprehensive review of programme specifications, unit specifications and assessment guidelines.

To complement this analysis, student perspectives were gathered through an anonymous questionnaire featuring both closed and open-ended questions. The questionnaire aimed to assess students' experiences with current assessment methods and identify areas for improvement. Responses were analysed using descriptive statistics for quantitative data, while qualitative feedback provided deeper insights into student concerns and expectations.

Stage three: outcomes

The analysis of assessment practices revealed several critical insights into the assessment pattern in the programme in terms of assessment load, methods, distribution and alignment with ILOs.

Load

Overall, the programme includes a total of 26 summative and five formative assessments, with a ratio of five summative to one formative assessment (Figure 2A). While the majority of modules employ three or four summative assessments, less than half of the modules incorporate formative activities (Figure 2B). This imbalance indicates that students are predominantly evaluated through summative assessments, which limits opportunities for

continuous, formative feedback and engaging in developmental, low-stakes learning experiences. These findings align with previous observations suggesting that research-intensive programmes tend to have a higher load of summative assessments (Tomas and Jessop, 2018).

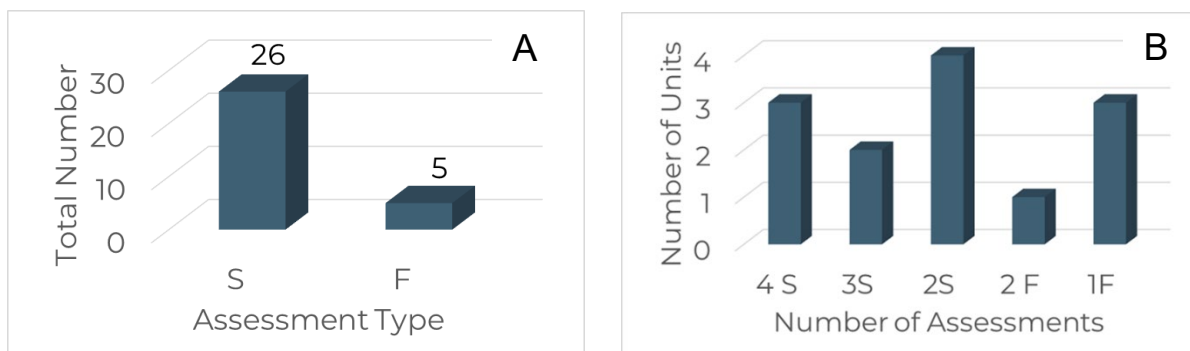


Figure 2A. Total number of summative (S) and formative (F) assessments. Figure 2B. Number of assessments per unit.

Methods

When mapping assessment methods across individual modules, a total of 10 different methods are used (Figure 3A). Of these, six are largely based on writing skills, including critical appraisals, proposals, short and long essays, reflections and the dissertation. Among all the methods, essays (both short and long) and multiple-choice questions are the most commonly used formats (Figure 3B). Essays feature prominently in two-thirds of the modules, while multiple-choice questions appear in more than half. Although these methods are valid, the over-reliance on these assessment types may limit students' opportunities to showcase a broader range of skills and disadvantage certain learning styles.

Diversifying assessments would more efficiently support student learning, improve student experience and engagement (Winstone and Carless, 2020; O'Neill and Padden, 2021).

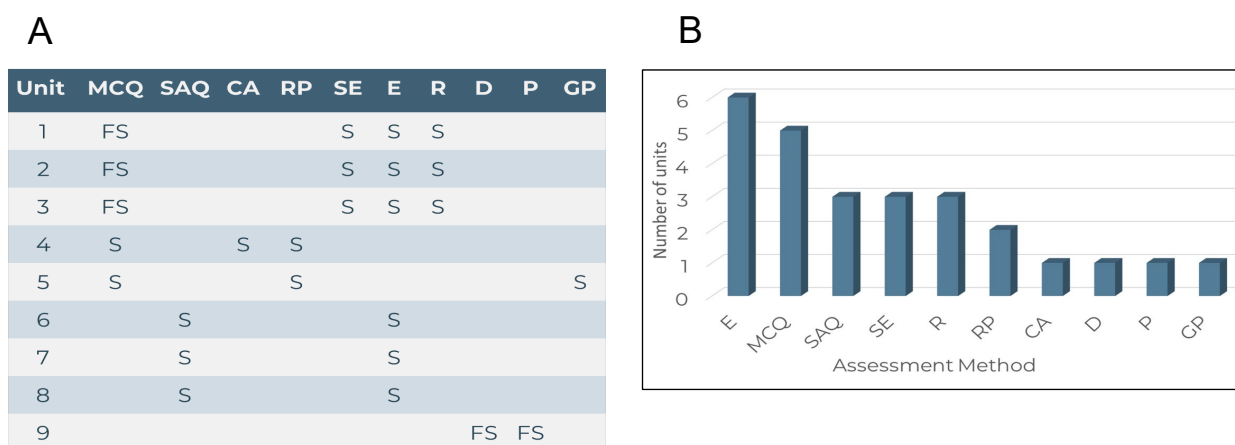


Figure 3A. Assessment mapping matrix. Figure 3B. Methods usage

E-Long Essay (2500 words); MCQ-Multiple Choice Question Exam; SAQ-Short Answer Question Exam; SE-Short Essay (1000 words); R-Reflective Piece; RP-Research Proposal; CA-Critical Appraisal; D-Dissertation; P-Presentation, GP-Group Presentation; F-Formative, S-Summative assessment.

Alignment with ILOs

In terms of alignment, the mapping exercise (Figure 4A) revealed that while assessments align with the programme ILOs, significant inconsistencies exist. Approximately half of the ILOs are assessed in just one or two assessments, while more than a quarter are over-assessed, being evaluated across four or more modules (Figure 4B). Furthermore, around 9% of ILOs are not assessed at all. This indicates a misalignment between assessment distribution and the broader learning goals of the programme, where certain areas of learning are overemphasised, and others are neglected (Bloxham and Boyd, 2007).

Distribution

The distribution of assessments throughout the academic year was found to be uneven, with 62% of all assessments concentrated in two periods: November/December and April/May (Figure 5). During these periods, some weeks see a load of up to three or four summative assessments. This clustering of assessments creates the risk of overload, leading to increased stress and anxiety among students. Furthermore, it may prompt a shift toward

strategic learning approaches, focused on grades and meeting deadlines rather than fostering deep, meaningful engagement with the material (Tomas and Jessop, 2018).

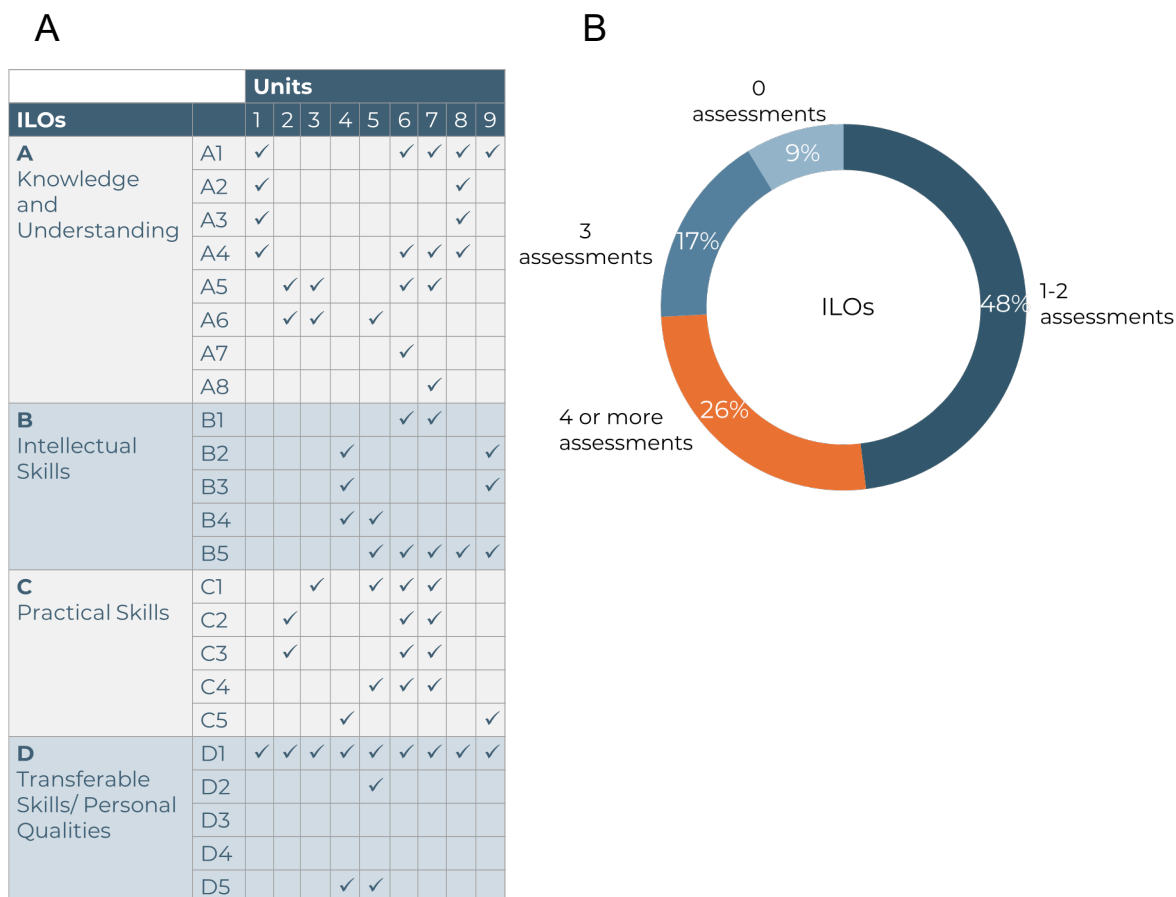


Figure 4A. ILO mapping matrix.

Figure 4B. Analysis of the frequency with which each ILO is assessed

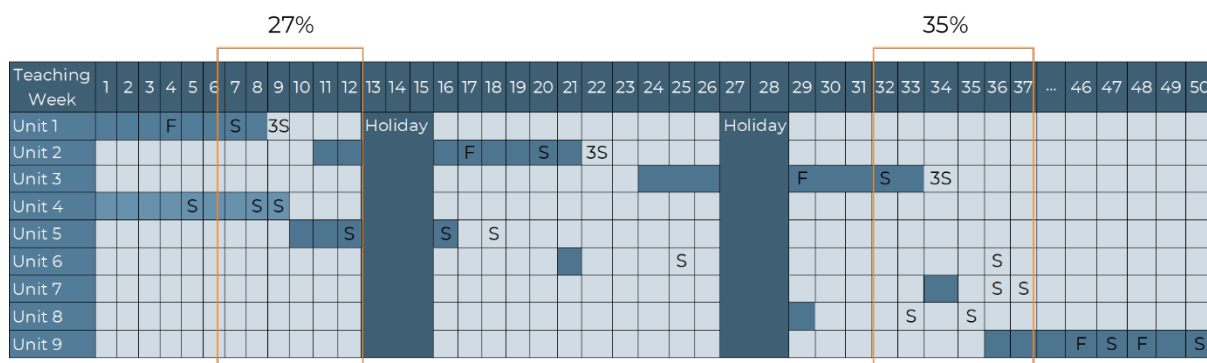


Figure 5. Assessment distribution across the academic year

Student feedback

Feedback collected from students regarding the inclusivity and authenticity of the assessments revealed mixed results. A significant majority, 80% (12/15), reported that the assessments were inclusive and aligned with their individual learning preferences. However, only 60% (9/15) felt that the assessments effectively represented real-world applications of the knowledge and skills they had acquired. All students expressed concerns about the excessive number of summative assessments within the programme.

When asked to identify the primary drawbacks, responses mainly focused on three key issues: assessment overload (80%, 12/15), clustering of assessment deadlines within tight timeframes (13%, 2/15), and delays in receiving feedback (7%, 1/15).

In terms of potential improvements, students suggested: reducing the overall assessment burden, redistributing assessments more evenly throughout the academic year to prevent overlapping deadlines, diversifying assessment types, and increasing opportunities for practical, problem-solving, collaborative and peer assessments with options for student choice. Additionally, students called for improved quality and timeliness of feedback, as well as clearer assessment instructions and marking criteria.

Action plan

The key output from the assessment audit and student questionnaire was a 10-point action plan outlining the key enhancements (Table 1). This action plan will inform the assessment enhancement process.

	Action	Benefit
1	Reduce the number of summative assessments	Reduced stress and anxiety, better engagement and quality of work
2	Increase the number of formative opportunities	Continuous feedback and engagement in developmental, low-stakes learning experiences
3	Reduce assessment bunching	Better time management, reduced stress and anxiety, better outcomes
4	Diversify assessments methods	More inclusive, catering to diverse learning styles, enabling students to showcase their knowledge and skills in ways that align with their strengths
5	Ensure balanced assessment of ILOs	Development of broad range of skills, more equitable evaluation of student performance

6	Provide choice in assessments	Fostering autonomy and ownership over academic progress, inclusive of various learning styles
7	Foster collaborative and peer-assessments	Development of collaboration and teamwork skills, enhanced critical thinking and reflection
8	Improve feedback quality and timeliness	Students better supported, better understanding of strengths and areas for improvement, better progress
9	Enhance assessment authenticity	More relevant learning experience connecting academic content with practical applications, developing key skills relevant to professional context, increased motivation and engagement
10	Assess the impact of any proposed changes to individual module assessments at the programme level before implementation	Consistent and aligned assessment practices, ensuring academic standards and quality of the programme

Table 1. Action plan for enhancing assessment practices

Key messages and transferability

Although assessment practices appeared relevant, effective and well-aligned at the individual module level, a broader programme-wide analysis revealed inconsistencies and areas for improvement. These insights informed the development of a structured, data-driven strategy for assessment redesign.

While this case study focuses on a postgraduate taught programme, the approach and methodology are adaptable across disciplines and academic levels, providing valuable guidance for programme leaders seeking to create an assessment environment that supports student learning and enhances the overall learning experience.

Next steps

As of now, the action plan is under review by the programme's academic team, focusing on developing appropriate assessment solutions at the module level and ensuring alignment across the entire programme. In addition, a student focus group has been formed to provide input into the co-design and decision-making process for assessments. This approach ensures that the redesign incorporates not only pedagogical principles and educational outcomes but also student needs and preferences. On completion of the redesign, the changes will be implemented and evaluated.

References

- Biggs, J (2003) *Teaching for quality learning at university*. Buckingham: Society for Research into Higher Education and the Open University Press.
- Biggs, J and Tang, C (2011) *Teaching for Quality Learning at University*. Maidenhead: McGraw-Hill Education (UK).
- Bloxham, S and Boyd, P (2007) *Developing Effective Assessment in Higher Education: A Practical Guide*. Maidenhead: McGraw-Hill Education.
- Boud, D. (1995). Assessment and learning: contradictory or complementary? In Knight, P. (Ed) *Assessment for learning in higher education* (1st ed.). Routledge, p. 35.
<https://doi.org/10.4324/9780203062074>
- Gibbs, G and Dunbar-Goddet, H (2009) 'Characterising programme-level assessment environments that support learning', *Assessment and Evaluation in Higher Education*, 34 (4): 481-489. <https://doi.org/10.1080/02602930802071114>
- Gibbs, G and Simpson, C (2004) 'Conditions under which assessment supports students' learning', *Learning and Teaching in Higher Education*, 1 (1): 3-31.
- Jessop, T and Tomas, C (2017) 'The implications of programme assessment patterns for student learning', *Assessment and Evaluation in Higher Education*, 42 (6): 990-999.
<https://doi.org/10.1080/02602938.2016.1217501>
- Jessop, T, El Hakim, Y and Gibbs, G (2014) 'The whole is greater than the sum of its parts: a large-scale study of students' learning in response to different assessment patterns', *Assessment and Evaluation in Higher Education*, 39 (1): 73-88.
- O'Neill, G and Padden, L (2021) 'Diversifying assessment methods: Barriers, benefits and enablers', *Innovations in Education and Teaching International*, 59 (4): 398-409.
<https://doi.org/10.1080/14703297.2021.1880462>
- Tomas, C., & Jessop, T. (2018). Struggling and juggling: a comparison of student assessment loads across research and teaching-intensive universities. *Assessment & Evaluation in Higher Education*, 44(1), 1–10.
<https://doi.org/10.1080/02602938.2018.1463355>
- Winstone, N and Carless, D (2020) *Designing Effective Feedback Processes in Higher Education: A Learning-Focused Approach*. London: Routledge.
- Wu, Q and Jessop, T (2018) 'Formative assessment: Missing in action in both research-intensive and teaching focused universities', *Assessment and Evaluation in Higher Education*, 43 (7): 1019-1031. <https://doi.org/10.1080/02602938.2018.1426097>
-

The how and WHEN of a successful feedback strategy

Giulia Getti, Associate Professor, University of Greenwich, UK

Stage one: background

The importance of effective feedback in higher education can not be overestimated. It enhances students' learning, encourages communication and improve outcomes and behavior (Hattie, 2007 and 2009; King, Schrodtt and Weisel, 2009; Poulos and Mahony, 2008; Juwah et al, 2004). Therefore, an effective feedback strategy is a great opportunity to positively impact on the quality of the students' learning process. Nevertheless, assessment and feedback (A&F) have consistently been challenging area in UK higher education institutions (Buckley, 2021), with assessment-related survey questions scoring low on the National Student Survey (NSS). At the same time, colleagues within the department have been sharing similarly negative experiences on the response of students to feedback, reporting that learners showed a consistent lack of interest in the feedback received, and failed to apply it. Dissatisfaction with assessment outcomes is linked to several factors, as clearly identified by Winstone et al (2017). Some of these factors require an improvement in clarity and accessibility of our feedback and instructions as well as the provision of actionable feedback. But this is not enough. Effective feedback can only be achieved by designing assessment in a manner that includes students in the learning process.

"No matter how quick, how detailed, or how high-quality the feedback our students receive, feedback can never be effective unless they use it."

(Nash and Winstone, 2017)

Therefore, the question this case study aims to look at is how to develop an assessment and feedback model that encourages learners to use the feedback received, stimulates dialogue, supports self-reflection and focuses on rewarding improvement, rather than just grade.

Stage two: approach

This case study refers to changes made to a final year core module for the Biomedical Science degree, which previously included two summative assessments, one essay and a phase test. The feedback on the assessments was rarely collected and there was no indication that it had been understood or applied to assessments.

The essay was replaced with a novel assessment, designed to enhance opportunities for feedback in a range of ways, and ultimately improve students' engagement with, and understanding of, feedback.

Science students are asked to work in groups of five to develop an (individually marked) critical review of an infectious pathogen. Learners are given the freedom to chose which aspect of the topic to cover (for example, epidemiology, clinical aspects, pathogenicity etc) and effectively work together to produce a scientific review for publication. Although each

individual is responsible for covering one aspect, the group must collaborate, as the final product will have to be consistently presented, and the topics must link to each other's while avoiding repetitions.

Students are invited to submit a draft of their summative assessment ahead of final submission. Rather than marks, the formative submission receives timely and detailed feedback from the tutor, which is immediately applicable to the summative task. The effectiveness of feedback is maximised by providing learners with guidance precisely when they need it most (Hattie and Timperley, 2007). This approach aims to maximise the interest in the feedback as well as its understanding. Once feedback is released, an interactive tutorial is organised to allow for further discussion to take place in class. Additionally, one-on-one discussions are encouraged for students who wish to receive any clarification on the comments provided.

At the same time, students are asked to provide peer feedback to their group member's submissions. This is based on the knowledge that the involvement of learners in activities that replicate the experiences of markers supports their understanding of the assessment standards (Sadler, 1989, 2009, 2010 and 2014; Bloxham and Price, 2015; Butler Shay, 2004). These approaches aim at reinforcing students' grasp of feedback, as many students struggle to understand feedback comments and interpret them correctly (Higgins, 2000). Both these exercises have seen feedback separated by the grade, so that by removing the focus on grades, learners can prioritise constructive comments and improvement.

Within this structure, students continuously engage in receiving, providing and discussing feedback. For the summative submission, learners must reflect on the feedback they have received, highlighting how they have applied it or explaining why they have not, as well as responding to comments. This authentic assessment method mirrors the submission process of a peer-reviewed article, offering students the opportunity to practice a real-world scenario.

The marking scheme for this assessment is constructed to shift the focus from merely judging the submission (and by extension, the student) to serving as a tool for supporting learning and guiding improvement. Students receive grades for providing feedback, as well as for their reflection and application of the feedback received. This approach allows tutors to recognise and reward both engagement and improvement in the final mark. As we use assessments and marks to evaluate students' performance, this approach sends a clear message to students, which is that their voice and participation in the process is an important part of learning.

Moreover, students are effectively given a say in the assessment and feedback process; they have an opportunity to discuss misunderstandings or concerns and ask questions as part of the assessment, therefore without feeling singled out. By being involved in the feedback cycle, their learning experience remains a positive and inclusive one.

Stage three: outcomes

The main change, which introduced a group course work element and the opportunity to submit a draft for early feedback, was introduced in the academic year 2011-12. This resulted in an immediate improvement in students' performance, with an average mark in the module increasing from 44 and 52 in 2008-9 and 2009-10 respectively to 64. As engagement with draft submission increased, so did students' performance, with an average mark of 70 consistently achieved in the following years.

Although students were provided feedback on drafts, it remained unclear whether the feedback was understood or applied. To address this, a second change was introduced, which consisted of including in the marking scheme 'how well feedback was applied'. By associating a proportion of the mark with engaging with the feedback a culture shift is prompted, where learning and improvement becomes focus of assessment and therefore rewarded. Although this change did not further affect performance, it did have an impact on students' response to course evaluation surveys.

The University has been collecting students' feedback through Evasys every year since 2014 and the survey consists of 16 questions separated into four themes, one of which is 'assessment and feedback'. Within this theme, two questions are specific to feedback, one asking whether feedback has been timely and the other, more relevant to this work, asking whether feedback (informal and formal) has helped students to develop and improve their performance (Q3.2). As providing feedback ahead of final submission aimed to support students' improvement, the answers to this question over the year are reported as representative of the impact of this feedback model on the cohort.

Student responses were assessed using five-point Likert scales, the average scored in Q3.2 for each year is shown in Figure 1. A clear increase in students' perception on whether the feedback has helped in improving their performance is visible in the main graph and summarised on the top left, where the average score before and after the introduction of feedback on marking scheme is shown.

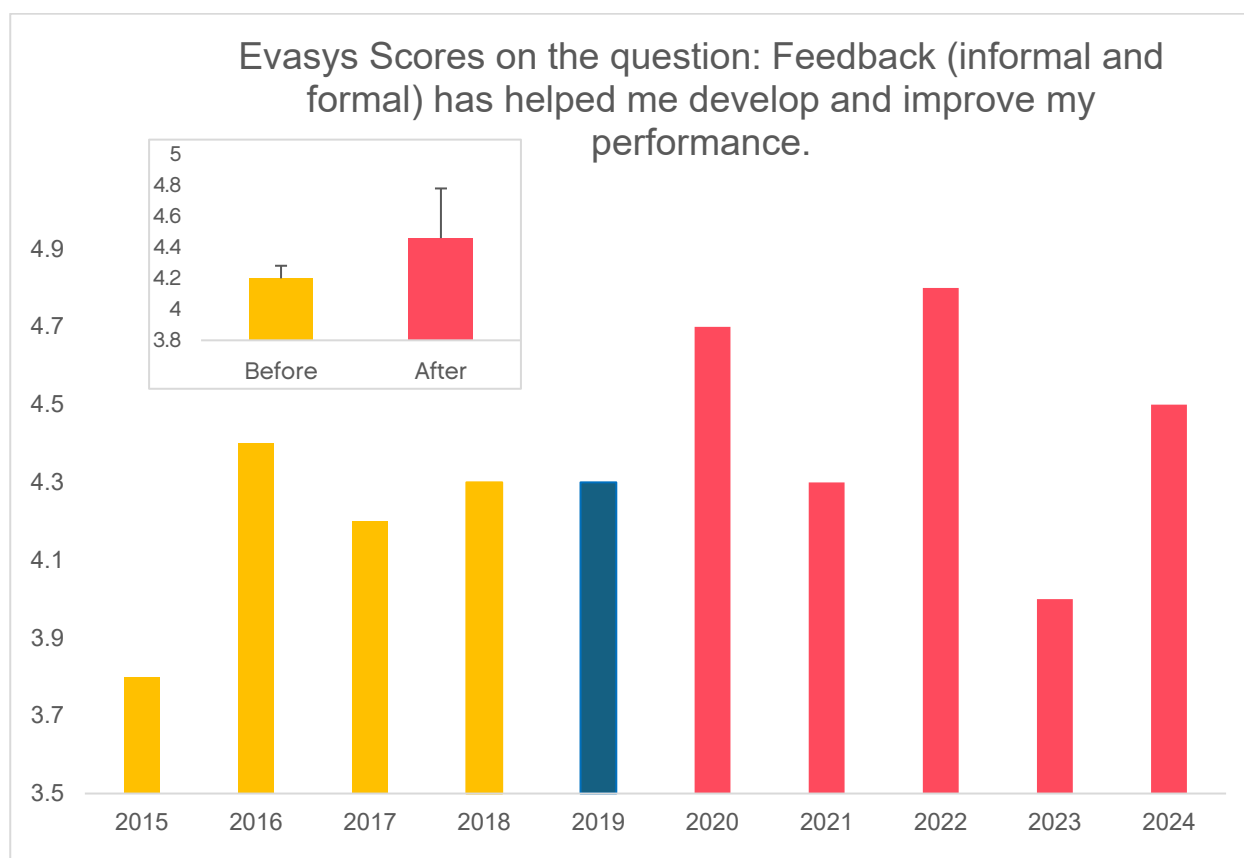


Figure 1. Evasys average scores for Q3.2 before (yellow) and after (blue) ‘how well feedback was applied’ has been introduced to marking scheme. The average scores over the years have been compared on the top left of the graph.

The fluctuation on the scores over the years is due to cohort effect (Nickerson, 2023) and challenges whether the improvement can genuinely be linked to the changes made in the feedback approaches. To further explore this point, data from a different module (Module 2) were collected and compared with data from the case study module (Module 1) (Figure 2). This second module was chosen because it runs at the same time as Module 1 and is made up of a similar cohort of students; both are core final year modules for the same programme. Module 2 uses standard feedback approaches, where learners receive mark and feedback on the final submission only.

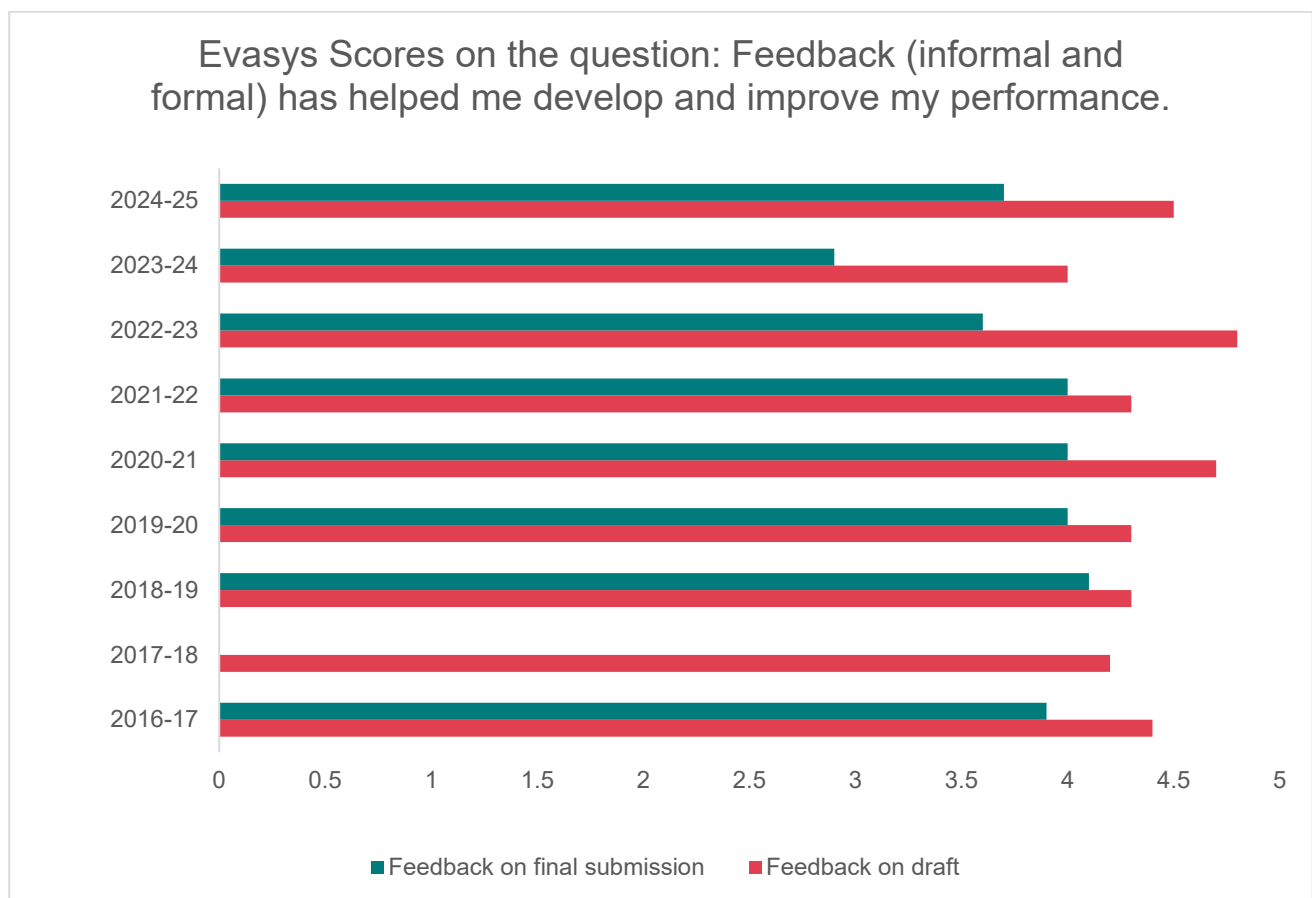


Figure 2. Evasys average scores for Q3.2 in Module 2, which uses standard feedback approaches, where learners receive mark and feedback on the final submission (in grey) and Medical Microbiology, where feedback is provided at draft stage (blue)

This latest overview clearly shows that receiving feedback ahead of final submission has had a consistently positive impact in students' perception of its helpfulness. Interestingly, the two cohorts where the effect of early feedback has been the highest (Figure 3), are the two with the highest and lowest average score in Module 1 on Q3.2 (2022-24: 4.8 and 2023-24:4). Those years I also experienced the overall highest and lowest level of student's engagement with the course. It will be interesting to carry out a more in-depth investigation to determine the effect of individual engagement on perception of feedback usefulness. Can it motivate the least engaged students?

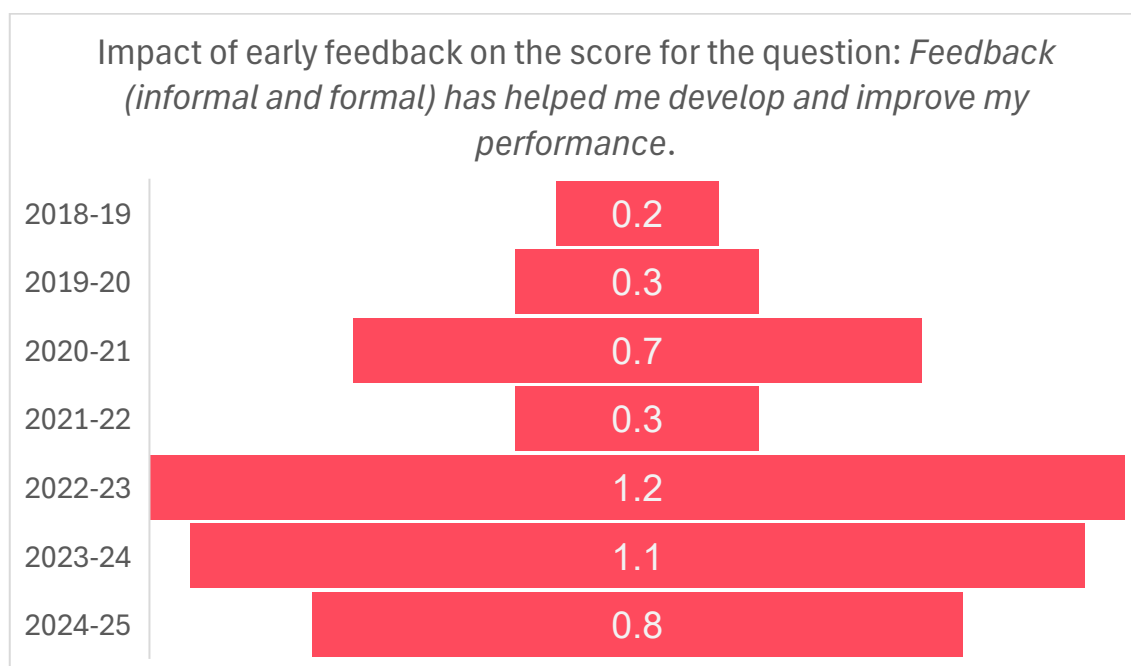


Figure 3. Difference in Evasys average scores for Q3.2 between Module 2, which uses standard feedback approaches, where learners receive mark and feedback on the final submission and Medical Microbiology, where feedback is provided at draft stage. All values indicate an increase in score when early feedback is provided.

Further evidence that this approach has had a positive impact was found in the feedback from the students themselves, which identified that the opportunity to submit a draft and receive feedback on it gave them the opportunity to apply the feedback before the final submission, and that this two-way communication between tutors and students is how they understand how to improve.

Assessment and feedback can do more than supporting topic-related learning and skills. A carefully designed assessment encourages all students to reflect on their learning, to share their thoughts and opinion, ultimately to be an active part of the learning process (HEA, 2013). Further investigation will be necessary to establish whether the above has been achieved with this model, but encouraging signs are present.

Encouraging a dialogue with students not only supports their learning but also allows the tutor to learn valuable lessons. For example, it has become clear that clarity of feedback means different things to academics and students; marks can function as a barrier to engaging with feedback; and feedback can negatively impact learners who lack confidence and see it as criticism. Therefore, using feedback as an opportunity to encourage and reassure students in their ability is a powerful instrument to maximise its outcome. The tutor benefits from engaging in these discussions as much as the students do, gaining insights into effective feedback practices and student perspectives.

References

- Bloxham, S and Price, M (2015) 'External Examining: Fit for Purpose?', *Studies in Higher Education*, 40 (2): 195-211.
- Buckley, A (2021) 'Crisis? What crisis? Interpreting student feedback on assessment', *Assessment and Evaluation in Higher Education*, 46 (7): 1008-1019.
- Butler Shay, S (2004) 'The Assessment of Complex Performance: A Socially Situated Interpretive Act', *Harvard Educational Review*, 74 (3): 307-329.
- Hattie, J (2009) *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hattie, J and Timperley, H (2007) 'The power of feedback', *Review of Educational Research*, 77 (1): 81-112.
- Higher Education Academy (2013) *HEA Feedback toolkit*. Available at: <https://www.advance-he.ac.uk/knowledge-hub/hea-feedback-toolkit> [accessed 10 December 2024].
- Higgins, R (2000) "'Be more critical!': Rethinking Assessment Feedback", in *British Educational Research Association Conference*, 7-10 September 2000, Cardiff University, Cardiff, UK.
- Juwah, C et al (2004) 'Enhancing student learning through effective formative feedback'. *Advance HE* Available at: <https://www.advance-he.ac.uk/knowledge-hub/enhancing-student-learning-through-effective-formative-feedback> [accessed 20 October 2025]
- Nash, R A and Winstone, N E (2017) 'Responsibility-sharing in the giving and receiving of assessment feedback', *Frontiers in Psychology*, 8, Article 1519. <https://doi.org/10.3389/fpsyg.2017.01519>
- Nickerson, C (2023) *What Is a Cohort Effect? Definition and Examples*. Simply Psychology. Available at: <https://www.simplypsychology.org/cohort-effect-definition.html> [accessed December 2024].
- Poulos, A and Mahony, M J (2008) 'Effectiveness of feedback: the students' perspective', *Assessment and Evaluation in Higher Education*, 33 (2): 143-154. <https://doi.org/10.1080/02602930601127869>
- Sadler, D R (1989) 'Formative Assessment and the Design of Instructional Systems', *Instructional Science*, 18 (2): 119-144.
- Sadler, D R (2009) 'Transforming Holistic Assessment and Grading into a Vehicle for Complex Learning', in Joughin, G (ed.) *Assessment, Learning and Judgement in Higher Education*. Dordrecht: Springer, pp. 45-63.
- Sadler, D R (2010) 'Beyond Feedback: Developing Student Capability in Complex Appraisal', *Assessment and Evaluation in Higher Education*, 35 (5): 535-550.
-

Sadler, D R (2014) 'The Futility of Attempting to Codify Academic Achievement Standards', *Higher Education*, 67 (3): 273-288.

Winstone, N E, Nash, R A, Parker, M and Rowntree, J (2017) 'Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of recipience processes', *Educational Psychologist*, 52 (1): 17-37. <https://doi.org/10.1080/00461520.2016.1207538>

Enhancing assessment and feedback in Master's projects through reflective logs and focus on project skills

Abdul Wadood Al-Ahad Sharif, Katarzyna Sypek, & Mark Haw, Department of Chemical and Process Engineering, University of Strathclyde, Glasgow, UK

Stage one: background

Master's 'dissertation projects' – the need for redesign

In this case study, we address two challenges with the traditional approach to assessment of the 'dissertation project', a key part of Master's-level curricula across many disciplines.

First, while such projects provide students with the opportunity to develop a wealth of deeper learning and skills (such as research, reflection, time management, critical thinking and open-ended problem solving) highly relevant to employability and wider personal development, assessment is often focused around the report or output text, which does not obviously provide students with either motivation to focus on wider learning, nor reward for achievement.

Second, the rise of easy-to-access generative AI means such reports are highly vulnerable to AI-assisted production, raising significant concerns around academic integrity (Overono and Ditta, 2023). Trying to deal with the use of AI (typically difficult to conclusively detect) through, for example, a plagiarism policy, opens major strategy challenges for HE institutions.

Our response to these issues aligns with several relevant themes including authentic assessment (Ajjawi et al, 2024; Maxwell, 2012), enhanced feedback practice, diversified assessment and academic integrity. Our approach is to refocus both the operation and the assessment of the 'dissertation' project on its basic purpose: to enable the student to develop deeper skills, knowledge and understanding and to demonstrate these clearly through assessment.

Context of case study

This case study explores redesign and implementation of changes in both operating and assessing a 60-credit individual project carried out by students registered on the MSc Advanced Chemical Engineering and Sustainable Engineering: Chemical Engineering programmes at the University of Strathclyde, Glasgow, Scotland. Typical student numbers are 20-40, demographically mixed as regards gender, nationality and ethnicity. The 60-credit project has wide commonalities across and beyond STEM subjects, demanding skills such as research, communication and presentation, innovation and problem-solving, contextualisation of the project topic, resource and time-management,

ethical considerations and reflection. Hence, the case study has significant relevance across disciplines despite its focus on this implementation in Chemical Engineering.

The redesign of project operations and assessment focuses attention (of both teachers and students) on how the student develops and uses key skills throughout the project, and how they reflect on their experience during and after the project (Fullana et al, 2016). This 'constant attention', contrasted with previous assessment dominated by the final report, aims to cement learning through experience as the student carries out the project, as well as providing a portfolio of evidence to underpin authentic final assessment.

Stage two: approach

The redesign involved a team comprised of the project coordinator, departmental assessment coordinator and director of PGT programmes, all of whom had significant experience with the projects. The entire process took approximately 10 months (six months planning, three months implementation, one month review).

Diversifying assessment – authentic assessment

Assessment weightings were rebalanced to emphasise project skills development, through the introduction of a performance and skills component (the process) with equal weighting of the report (the written product) and a separate presentation and Q&A component.

Original assessment	Redesigned assessment
12,000 word report (80%)	7,000 article submission (40%)
Project Supervisor Assessment (20%, including presentation)	Continuous Performance & Skills (40%)
	Presentation and Q&A (20%)

To improve students' engagement with the revamped performance and skills assessment, students were required to submit regular reflective logs over the duration of the projects, each followed by a formative reflective meeting for verbal and brief written feedback provided by supervisors (Robinson and Udall, 2006). Language around assessment was deemed important – note the subtle shift in emphasis from the supervisor (source of assessment) to the capabilities of the student (performance and skills), which included technical understanding; problem-solving; planning; communication and reflection. These logs were structured with reference to the assessment criteria.

The written submission, originally a 12,000-word report, was reduced to a 7,000-word journal article submission, with greater emphasis placed on the research methodology, design, analysis and coherence of the piece; addressing possible GenAI vulnerabilities. Students were also provided with an opportunity to submit a draft report with feedback from their report assessor.

Workload management – supervisory pairings

Each project involved supervisory pairs who worked together to develop a project proposal, supervise students and moderate one another's assessment of performance and skills. This introduction resolved a few challenges that had emerged around project delivery:

- + facilitation of annual leave – academics take leave during summer and so pairings ensure continuous supervision for students throughout the projects
- + scalable projects – project proposals were designed to facilitate a variable number of students (academic year 2021-22 observed a two-fold increase in cohort size) and allow for fair distribution of supervision workload
- + quality assurance around assessment – each student was allocated a 'primary' supervisor from the pair responsible for performance and skills assessment which could be moderated by the 'secondary' supervisor.

Consistent feedback – marking rubrics

There were also challenges around high variability in volume, language and tone of feedback, as well as the potential subjective interpretations for high-level criteria from assessors.

To address this, for each assessment a detailed criteria and associated sub-criteria were developed with generalised feedback comments and rubrics developed for each sub-criteria aligned to the University's descriptors for feedback (outstanding, satisfactory etc). The sub-criteria were developed to ensure that each criterion was assessed holistically. To address shortfalls of generalised feedback from rubrics, specific feedback was also expected with guidance (such as approximate word count and balancing comments on areas of merit/areas for improvement).

Facilitation and communication of change

A key enabler of this redesign was inviting those impacted to influence the changes. Several consultations with academics who were involved in supervision took place during the months prior to implementation. Feedback received from these consultations was incorporated into the final design.

Students were informed of the changes and presented with these at the induction session for the projects. The rationale and justification presented in this case study was communicated to students.

Alongside communicating the changes, workshops were introduced for both academics and students. The training workshops for academics demonstrated how the performance and skills reflective logs would function and defined roles and responsibilities, and another workshop communicated expectations around assessment and feedback. Students had several workshops that covered multiple areas including reflective and academic writing; research ethics; generative AI use for research; and presentation skills.

Stage three: outcomes

Upon completion of the project, both students and supervisors were asked to participate in separate surveys, one for students and another for supervisors, for which ethical approval was secured.

Satisfied students

Overall, students' outcomes improved with the median grade increasing from 61.7% from the year prior to 66.5%. Results indicate that students were generally satisfied with the various components of the project (Figure 1). Despite a low response rate from students (cohort of 25, n = 7), these were triangulated through informal comments from students with students' and supervisor's surveyed perspectives. The highest satisfaction was noted for staff support (communication, coordination, workshops and handbook), underscoring the importance of clear communication when implementing changes to such activities. Assessment-related items were also well received, with over 70% of respondents being very satisfied with the report format (journal article). The presentation and evaluation criteria were positively received as well. For reflective logs, 43% of students were very satisfied, 29% were neutral, and 29% were somewhat dissatisfied.

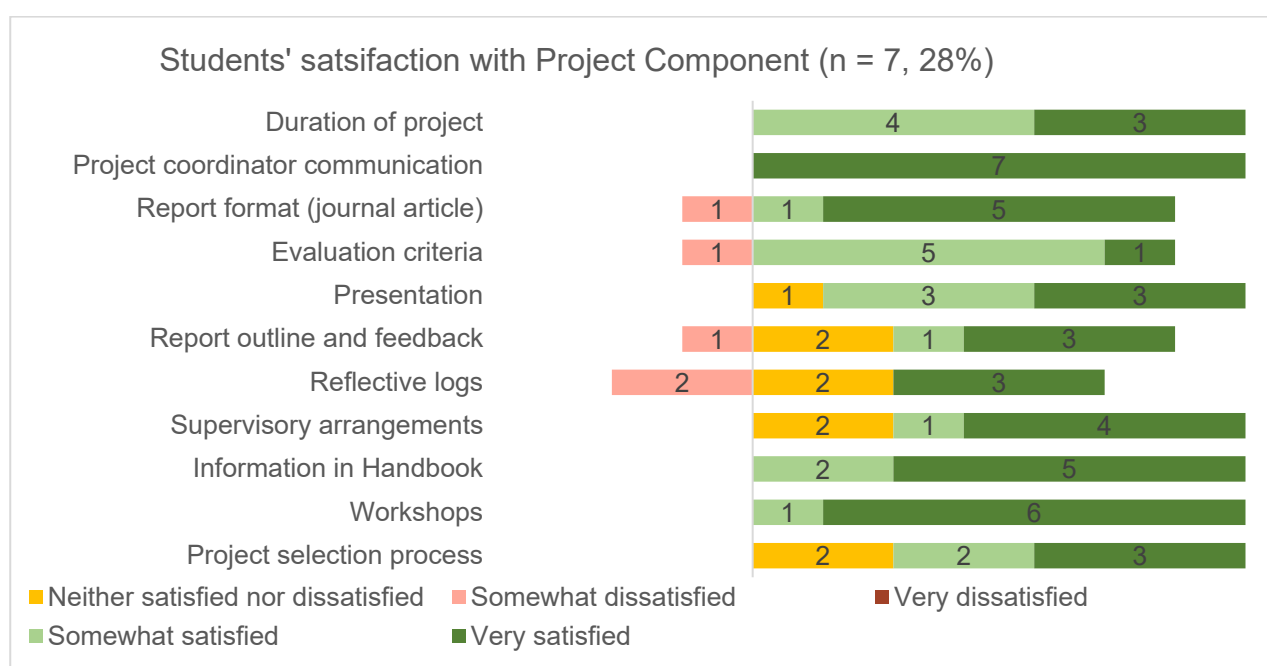


Figure. 5 Students' satisfaction with components of MSc projects

In the free comments on workshops, students mentioned they were informative and helpful but suggested these were at times difficult to apply to the specifics of a given project. Comments on reflective logs and performance and skills assessment indicated that students found them useful for tracking their progress. One student noted,

“I think it is a creative way to track the progress of the project, which was useful for us as well.”

However, one student stated there was the view that the frequency of logs (fortnightly) resulted in repetition. Another student mentioned they struggled with writing certain sections, such as ethics, collaboration and trust; indicating more guidance is necessary here.

Satisfied supervisors

The supervisor survey results were very positive overall (Figure 2). Most supervisors were satisfied with the communication from the project coordinator, supervisory arrangements, evaluation process and evaluation criteria (rubrics). The implementation of reflective logs was well received, with 17% very satisfied, 50% somewhat satisfied, 17% neutral, and 17% somewhat dissatisfied.

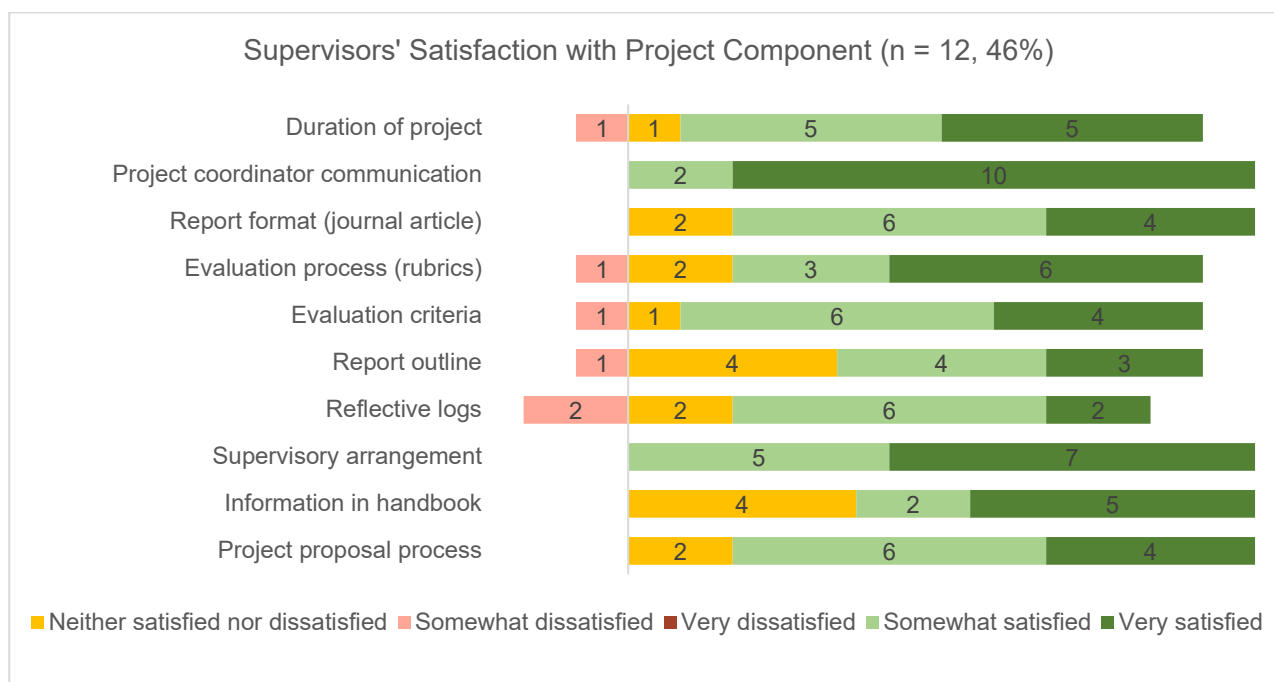


Figure 6. Supervisors' satisfaction with components of MSc projects

Free comments on reflective logs and performance and skills assessment indicated that while staff appreciated the usefulness of the logs, they would prefer fewer of them, which mirrors comments from students. Staff observed increased student ownership of projects, improvement in writing, enhanced engagement and a tendency to view challenges as learning opportunities. One staff member commented:

“This was a useful exercise and made the students think more about what they were doing and why. I think it also engendered a greater sense of ownership for the project.”

Regarding assessment, supervisors appreciated the use of rubrics, although one noted,

“I feel that the process was too fine-grained; we should not slice and dice the reports so much.”

Supervisory arrangements in pairings were also well received, with comments such as, “I found having weekly meetings and working in pairs to be an excellent way of working,” and “Pairing supervisors allowed for some flexibility in terms of taking annual leave in summer.”

Planned improvements – next steps and reflections

Given the benefits associated with reflective logs, we will continue with these but reduce the number to ensure that both supervisors and students continue to appreciate the value of these and do not feel ‘assessment fatigue’. To address difficulties around ethics, collaboration and trust, we will encourage students to reflect on their own projects critically and expand their notion of ethics and trust in particular. To make the reflective logs more inclusive, we aim to promote student choice in how these logs are delivered and will accept audio recordings or vlogs.

It was noted that there was still an inclination for students to focus on lengthy literature reviews, despite being awarded limited marks. We believe that language plays an important role and may have contributed to this tendency. In the project information, the submission was still called a “report” despite the requirements and going forward we will adopt a more authentic name for the submission: “research article”.

We believe that the assessment design, which includes the Assessment of Skills and Performance, is valuable and can be effectively used for other major projects, such as design projects, Master’s projects, and undergraduate projects.

References

- Ajjawi, R, Tai, J, Dollinger, M, Dawson, P, Boud, D and Bearman, M (2024) 'From authentic assessment to authenticity in assessment: broadening perspectives', *Assessment and Evaluation in Higher Education*, 49 (4): 499-510.
- Fullana, J, Pallisera, M, Colomer, J, Fernández Peña, R and Pérez-Burriel, M (2016) 'Reflective learning in higher education: a qualitative study on students' perceptions', *Studies in Higher Education*, 41 (6): 1008-1022.
- Maxwell, T W (2012) 'Assessment in higher education in the professions: action research as an authentic assessment task', *Teaching in Higher Education*, 17 (6): 686-696.
- Overono, A. L., & Ditta, A. S. (2023). The Rise of Artificial Intelligence: A Clarion Call for Higher Education to Redefine Learning and Reimagine Assessment. *College Teaching*, 73(2), 123–126. <https://doi.org/10.1080/87567555.2023.2233653>
-

Robinson, A., & Udall, M. (2006). Using formative assessment to improve student learning through critical reflection. In C. Bryan & K. Clegg (Eds.), *Innovative assessment in higher education* (1st Ed. pp. 92-99). Routledge.

Supporting students with resit assessments

**Karen Lander, Senior Lecturer; Caitlin Hunter and Ruby Prior, Student Partners
University of Manchester, UK**

Stage one: background

In this case study, we explore the issues experienced and the strategies used at the University of Manchester for supporting students with resit assessments. The University of Manchester is a large University (44,000+ students), committed to world-class research, an outstanding learning and student experience and social responsibility. This project involved students as partners, with data collected via student focus groups.

Resits provide “a second chance” (Ricketts, 2010, 351) for students when they have not passed a component of their course. At the University of Manchester, a student is deemed to have ‘not passed’ a summative assessment if they are awarded a mark below 40% (undergraduate degree). A resit requires the student to achieve a mark $\geq 40\%$, to demonstrate achievement of the unit intended learning outcomes (ILOs) and be awarded the associated unit credits, as required for progression on the course (Juriansz, Nguyen and Clark, 2022). While some students may be required to resit, others are instead be able to ‘compensate’ marks between 30-40%. Here, students compensate insufficient performance on one unit with better performance overall (Haladyna and Hess, 1999; Schmidt, Cohen-Schotanus and Arends, 2021).

In general, the number of HE students needing to resit assessments is significant (estimated to be between 5% - 25%), varying by university, course, level and unit (Hunt and Chalmers, 2021). While we want to support and engage our students to make the most of this ‘second chance’, resits often pose a significant challenge to them. Indeed, resits provide a barrier to progression and typically create significant stress and disruption (Nguyen, Clark and Juriansz, 2023). At the University of Manchester, the resit period is in August – a time where many students are working or are away, making them somewhat detached and disengaged from the academic environment (Smith and Brown, 2020). In addition, having failed the original assessments, students with resits may lack academic confidence (Farrant et al, 2021). They may find it difficult to admit failure and consequently may not seek advice on how to turn failure into success (Jones and Smith, 2022).

While Ricketts (2010) proposes there is ‘no theory of resits’, we use Bandura’s (1999) social cognitive theory to draw attention to the importance of self-belief and efficacy on an individual’s learning and development. Here, students are themselves proactive agents, influencing their academic performance through self-reflection and confidence to act (Geertshuis, Jung and Cooper-Thomas, 2014). Learning is therefore not dictated only by university (externally) but also internally through self-efficacy and belief. Indeed “there is evidence (Bandura, 1997) that self-efficacious students... have fewer adverse emotional reactions when they encounter difficulties than do those who doubt their capabilities”

(Zimmerman, 2000, 86). Self-efficacy is an important skill for students to develop and this is especially the case when undertaking resits (Farrant et al, 2021).

Overall, the aims of this project were to (i) to explore the issues experienced (or the issues student think they would experience) when preparing for resits and during the resit period; and (ii) to consider the usefulness of external and internal support strategies to students, highlighting evidence of good practice.

Stage two: approach

In this project, a collaborative approach was used with input from an academic member of staff (Karen Lander) and a team of student partners (led by Caitlin Hunter and Ruby Prior). This project took place over the 2023-24 academic year, with a total two-year period needed from project conception, granting of ethical approval, data collection, thematic analysis and writing up. Overall, we conducted a series of focus groups (one hour) with students from across the University of Manchester. Students from courses based within two faculties contributed – Faculty of Biology, Medicine and Health (Psychology) and Faculty of Science and Engineering (Chemistry).

The focus groups included activities intended to encourage discussion among student attendees (see Colucci, 2007). A total of 12 focus groups were conducted (up to 10 attendees per group). Braun and Clarke (2021) propose that when using reflexive thematic analysis, “how many” data items, and when to stop data collection, are “inescapably situated and subjective” (331). Thus, we took a pragmatic approach, choosing a feasible number of focus groups that would allow data to be collected and analysed within the time constraints of the project.

In each focus group, one student partner led the session, and two further student partners took detailed notes on the discussions (labelling attendees by number). We were precluded from explicitly recording the focus groups and transcribing them due to ethical considerations at the University of Manchester. Student attendees who took part in the focus groups did not need to have direct experience of undertaking resits and, indeed, we explicitly told them not to disclose any of their own personal circumstances. Thus, student attendees could answer in relation to ‘hypothetical’ issues they could envisage and strategies that may help. In the focus groups we fostered a discussion between attendees, using a series of semi-structured questions as follows:

- 1 If you needed to undertake a resit, what problems and issues might you encounter leading up to or during resit periods?
 - 2 Do you know of what support mechanisms already in place at the University of Manchester to help support students during resits? If so, what are they and how do they work?
 - 3 What kind of support would you ideally want from the University when preparing for a resit?
 - 4 Do you think there are any strategies you could use to help yourself improve your academic confidence during resits?
-

After a general discussion surrounding these questions, attendees were given a list of proposed support mechanisms and asked to discuss them within the focus group and rank order them in terms of their usefulness. The support mechanisms and corresponding rankings are listed in Table 1. The strategies were created by gathering opinions from the student partners in conjunction with senior academic advisors from the University of Manchester. Student partners were given time to ask follow-up questions to allow free-flowing, meaningful conversation. Student attendees were sent vouchers after the focus group to thank them for their participation.

The detailed notes taken by the student partners during the focus group discussions were collated and we identified common themes through qualitative reflexive thematic analysis (Lochmiller, 2021). Here we focused on the issues students might encounter leading up to or during resit periods. Thematic analysis (Braun and Clarke, 2006, 79) involves “a method for identifying, analysing and reporting patterns (themes) within data”. The data was approached without prior theoretical assumptions and without reference to unarticulated meaning or social context, allowing the views and experiences of attendees to be understood from their own perspective (Braun et al, 2018).

Two researchers independently coded all the data, following the six-step technique advocated by Braun and Clarke (Braun et al, 2018), meeting regularly to discuss their interpretations. The notes taken were annotated to identify features (words, sentences or paragraphs) related to the scope of analysis and the associated ‘codes’ were organised into themes. Overall, three main themes were identified. Theme one involved issues surrounding ‘Reduced contact’, with sub-themes considering ‘Lack of contact and information from University’ and ‘Detachment from peers’. Specifically, they expressed that *“It would be really useful to have someone to talk to on my course about what went wrong last time”* (FG3, P5) and that they felt *“... out of touch and isolated”* (FG5, P1).

Theme 2 involved ‘Lack of unit specific knowledge’. Here, focus group members reported that one central issue regarding resits was poor knowledge of unit content and that *“If you don’t understand it to begin with then it is difficult to know how to improve”* (FG4, P6). A further example was *“I don’t get statistics. If I failed it first time then I wouldn’t know where to start to pass”* (FG12, P6). A final theme concerned ‘Academic confidence’, with students expressing that they were always going to *“do badly on some topics”* (FG1, P7) and they were *“not confident”* in knowing what to do.

As well as analysing the notes from the focus groups, results from the support strategy ranking task were collated and are shown in Table 1. Overall, the most useful strategy proposed was ‘Revision embedded within class time’, presumably with many students thinking that this would reduce the number of students not passing. Further, specific strategies surrounding content were deemed useful, such as generic or specific advice going over their previous exam. Interestingly, overall revision strategies were not thought useful, but instead attendees thought there may be some benefit in confidence boosting strategies, helping students to help themselves and perhaps tapping into improving self-belief.

Focus Group	1	2	3	4	5	6	7	8	9	10	11	12	Mean
'Revision class time' – set time for revision	2	1	3	1	4	3	1	1	2	1	1	4	2.00
Examples of previous exam questions with hints	1	4	2	3	1	2	3	4	1	3	3	2	2.42
Go over previous exam to demonstrate where they went wrong	3	5	1	4	2	1	5	2	4	2	2	1	2.67
Classes over the summer to provide catch up content on courses before resits	6	2	4	2	3	4	2	3	3	8	6	7	4.17
Information about confidence boosting strategies – like positive affirmation	5	3	5	5	6	5	4	5	6	4	4	3	4.58
Provide one person contact (Academic Advisor) at the University to speak to about resit	4	7	7	8	5	9	7	6	5	6	7	6	6.42
Link with a peer mentor for support	7	6	9	6	7	6	8	8	8	7	5	8	7.08
Set goals with 'exam mentor'	9	8	6	9	9	7	6	9	9	5	8	7	7.67
Access to study skills classes and 'how to revise effectively' tool kit	8	9	8	7	8	8	9	7	7	9	8	9	8.08

Table 1. Showing suggested support strategies along with the ranking of usefulness order by focus group and with overall mean. Strategies are listed in terms of usefulness with the most useful ranked first.

A limitation of this project is that we were unable to record the focus groups and thus the discussion was limited to notes rather than a transcription. Such notes may be subject to interpretation with some including more detail and some less. We tried to mitigate against this problem by assigning two student partners note-takers to each focus group. Future research would aim to record focus groups, and explicitly compare the views of students on different programmes and in different year groups. At the University of Manchester, the consequences of resits in year one are much less severe, with year one marks not contributing to the degree average.

Stage three: outcomes

Despite these limitations, our findings allow us to start to consider more carefully the issues and problems found by students in the periods leading up to and during resits. Our results suggest that students may experience (or anticipate) issues regarding reduced contact, lack of unit-specific knowledge and those related to academic confidence. We identified

examples of good practice, and our aim is to produce a set of guidelines about how universities can help effectively support students during resit periods.

Future work should also focus on also talking to students who did not pass their resits, as well as explicitly mapping the outlined strategies to the issues identified. For example, improving 'unit-specific knowledge' by offering both generic and tailored information about previous exam attempts. Finally, further work should explore the implementation of the strategies highlighted, measuring their direct impact on resit success (or failure) and academic achievement. Such research will contribute to this under-researched area and determine the best strategies (both external and internal) to help support future cohorts of students through the difficult time of resits.

References

- Bandura, A (1997) *Self-Efficacy: The Exercise of Control*. New York: W H Freeman.
- Bandura, A (1999) 'Social cognitive theory: An agentic perspective', *Asian Journal of Social Psychology*, 2: 21-41. <https://doi.org/10.1111/1467-839X.00024>
- Braun, V and Clarke, V (2006) 'Using thematic analysis in psychology', *Qualitative Research in Psychology*, 3 (2): 77-101.
- Braun, V and Clarke, V (2021) *Thematic Analysis: A Practical Guide*. London: SAGE Publications.
- Braun, V, Clarke, V, Hayfield, N and Terry, G (2018) 'Thematic Analysis', in Liamputtong, P (ed.) *Handbook of Research Methods in Health Social Sciences*. Singapore: Springer, pp 17-37.
- Colucci, E (2007) 'Focus groups can be fun: The use of activity-oriented questions in focus group discussions', *Qualitative Health Research*, 17 (10): 1422-1433. <https://doi.org/10.1177/1049732307308129>
- Farrant, C, Marinciu, R, Reilly, D and Warren, L (2021) 'Supporting students with resits to build self-efficacy', *Compass: Journal of Learning and Teaching*, 14 (2). <https://doi.org/10.21100/compass.v14i2.1185>
- Geertshuis, S, Jung, M and Cooper-Thomas, H (2014) 'Preparing Students for Higher Education: The Role of Proactivity', *International Journal of Teaching and Learning in Higher Education*, 26 (2): 157-169.
- Haladyna, T and Hess, R (1999) 'An Evaluation of Conjunctive and Compensatory Standard-Setting Strategies for Test Decisions', *Educational Assessment*, 6 (2): 129-153.
- Hunt, L and Chalmers, D (2021) *University Teaching in Focus: A Learning-centred Approach*. Abingdon: Routledge.
- Juriansz, J, Nguyen, N T and Clark, C (2022) 'Resit exams in Australian higher education: lessons from a novel COVID-19 assessment trial', in Singh, U G, Nair, C S, Blewett, C and
-

Shea, T (eds) *Academic Voices: A Conversation on New Approaches to Teaching and Learning in the Post-COVID World*, pp 429-441. <https://doi.org/10.1016/B978-0-323-91185-6.00034-3>

Lochmiller, C R (2021) 'Conducting Thematic Analysis in Qualitative Research: A Practical Guide', *Educational Researcher*, 50 (2): 123-133.

Nguyen, N T, Clark, C and Juriansz, J (2023) 'Effects of resit exams on student progression: An Australian case study', *Journal of University Teaching and Learning Practice*, 20 (1). <https://doi.org/10.53761/1.20.01.06>

Ricketts, C. (2010). A new look at resits: are they simply a second chance? *Assessment & Evaluation in Higher Education*, 35(4), 351–356. <https://doi.org/10.1080/02602931003763954>

Schmidt, H G, Cohen-Schotanus, J and Arends, L R (2021) 'Balancing Act: How Students Compensate for Poor Performance in Higher Education', *Assessment and Evaluation in Higher Education*, 46 (2): 234-248.

Terry, G, Hayfield, N, Clarke, V and Braun, V (2017) 'Thematic analysis', in Willig, C and Stainton Rogers, W (eds) *The SAGE Handbook of Qualitative Research in Psychology*. 2nd edn. London: SAGE Publications, pp. 17-37.

Zimmerman, B J (2000) 'Self-Efficacy: An Essential Motive to Learn', *Contemporary Educational Psychology*, 25 (1): 82-91.

Enhancing student feedback: a semi-automated approach using Microsoft Excel

Kyprianos Georgiou, Senior Lecturer in Forensic Science, University of West London; Sarah Hadden, Validation Technician & Lata Gautam, Associate Professor, Anglia Ruskin University, UK

Stage one: background

The Office for Students (2021) has reported a 4% decline (from 72.6% in 2020 to 68.59% in 2021) in UK students' satisfaction levels regarding assessment and feedback. This decline has continued into 2022, although this is a significantly smaller decrease of 0.06% (2022). In 2024, the satisfaction level increased to 78.3% (Office for Students, 2024). Even though this has increased from the previous year, it still falls behind satisfaction levels recorded for other vital student experience elements. For example, according to INTO University Partnerships (2024), UK students achieve satisfaction rates beyond 87% regarding their learning experience and reach satisfaction scores of up to 91% for support services, along with 95% for campus environment (international students). Additionally, the National Student Survey in 2024 revealed that certain departments and institutions recorded overall satisfaction rates above 80%, while specific subjects, such as Geography at Royal Holloway, reached 100% (according to Royal Holloway, 2024).

There are multiple reasons students could be feeling less satisfied with their assessment and feedback, including the disparity between student and staff expectations surrounding feedback (Dawson et al, 2018). Zhu, Liu and Lee (2019) have shown that students use feedback to review and adjust their work. This is supported by Dawson et al (2018), who showed that 49% of students stressed that feedback, including what to focus on, would be the most effective whereas only 12% of staff felt that including what to focus on was important, which indicates the disparity between expectations. This is further supported by Winter and Dye (2004), who found tutors felt that students perceived their grade as the feedback. Thus, tutors felt students were not willing to further interact with their provided feedback (Winter and Dye, 2004).

Staff time limitations may also offer explanations for the decline in student satisfaction. High student-to-staff ratios mean staff have limited amounts of time to prepare students' feedback (Barker, 2011) and higher staff workload. Zhu, Liu and Lee (2019) note the expectation of one teacher to provide all students in a single classroom feedback in a timely manner to be unrealistic. One method of improving time spent marking and providing feedback is the use of additional staff members. This alters the ratio of staff to students but may generate inconsistency among markers as some marking can be subjective (Nicoll, Douglas and Brinton, 2022).

There has been a movement towards more online spaces for learning in the education sector (Brinton and Chiang, 2014). This has been exacerbated by the coronavirus

pandemic, which required movement to online systems due to restrictions placed on travel/social distancing. There is also regular use of automated computation in tasks such as attendance and learning management systems, such as Canvas. With regards to feedback and marking, automation can be used within multiple capacities. Whether this be automated collation of marks, grading (Taghipour and Ng, 2016), or generation of feedback, which is the primary focus of this research.

Semi-automated marking and assessment systems are not a novel idea, with many studies offering their individual take on the creation of fully and semi-automated marking and feedback tools (Barker, 2011; Buyrukoglu, Batmaz and Lock, 2019; Hepplestone and Mather, 2007; Zhang and Hyland, 2018). However, many of these approaches tend to be faculty specific (Buyrukoglu, Batmaz and Lock, 2019; Calvo and Ellis, 2010; Zhang and Hyland, 2018). Semi-automated marking schemes save staff marking time and bring consistency, while still allowing personalised feedback for students (Barker, 2011; Buyrukoglu, Batmaz and Lock, 2019). This indicates the importance of developing, piloting and implementing semi-automated marking schemes, when possible, as proposed in this research. Thus, this study aims to detail the creation of our marking tool and to explore current literature surrounding semi-automated marking systems.

Stage two: approach

Initially the marking scheme used by the faculty team was reviewed through marking schemes collected from different colleagues and combined to see what could be improved. This was then developed further into the semi-automated tool using Microsoft Excel (Figure 1). The semi-automated tool proposed in this study is structured within four columns in Microsoft Excel:

- 1 The first column contains the beginning of a sentence intended to be the beginning of a sentence of a student's feedback (ie "The audibility during the presentation")
- 2 The second column contains letters a-e indicating options the user has (linked with the fourth column)
- 3 The third column contains empty cells, which is where the user can place an "x" next the letter (next column) indicating the ending of the sentence
- 4 The fourth column contains a series of statements intended for the various possibilities for the ending of the sentence (ie "was clear" or "needed improvement" (etc).

1) Poster presentation..... / 25 marks				Overall feedback	Total mark
Please put "x" in this column to reflect your feedback.					
The audibility during the presentation	a	x	was clear.	The audibility during the presentation was clear. The explanation was good. The pace was good. The presentation of points/flow was good. The student/presentation was enthusiastic and engaging.	25
	b		needs improvement.		
	c		was shaky and nervous.		
	d				
The explanation was	a	x	good.	Feedback	
	b		clear and concise.		
	c		confusing.		
	d		incorrect at times.		
The pace was	a	x	good.		
	b		either too fast or too slow.		
The presentation of points/flow	a	x	was good.		
	b		needs improvement.		
The student/presentation was	a	x	enthusiastic and engaging.		

Figure 7. This figure shows a screenshot of various criteria assessed using the semi-automated tool. At the bottom of the tool, there is also a free-text section whereby the marker can also provide personalised feedback for the student

This tool (among others) uses the function “=VLookup()”, which is designed to identify the users’ choices (the ending of the sentences) and collating them in a single cell displaying the overall feedback for the students’ presentation. A cell at the bottom of the marking scheme also gives the user the option to include free-text comments for more personalised feedback.

Articles were researched through databases such as Google Scholar, ARU Library and Science Direct. The following search phrases were used: “semi-automated student feedback generation”, “automated feedback tools for marking”, “automated student feedback”, “automated essay scoring”, “semi-automated marking feedback generation” and “maintain consistency in student marking use semi-automated”. Due to the mass amount of literature, relevancy was of key importance. Thus, articles were initially screened and then limited by relevancy. These were then filtered by year to only include those from 2007 onwards, with 23 articles being found and nine articles being selected for further comparison. These were sorted by relevancy, and chosen texts were organised into faculty and non-faculty specific as often these tools are related to specific disciplines. These texts were then evaluated and compared to our current semi-automated marking tool in development.

Stage three: outcomes

Our tool’s feedback is generated by pre-populating marking criteria as sentences with selectable choices for staff during marking (see Figure 1). It was proposed to use the tool for marking poster presentations of Level 6 students in BSc (Hons) Forensic Science and BSc (Hons) Crime and Investigative Studies courses. Students are marked based on criteria including presentation, understanding, layout and design, content, coverage and sources.

After two markers assessed 70 student posters, marks were higher on average for manual marking. A focus group was conducted to gather more feedback from colleagues.

A focus group consisting of 15 colleagues was organised from various subject areas. The group used the marking tool in a mock poster presentation, comparing it with manual marking. Most participants produced similar marks with both methods. Colleagues praised the tool as “very easy/simple to use” and “quick and easy.” However, one user struggled, noting a significant difference between manual and marks given by the semi-automated tool, indicating a learning curve and potential need for further development. The main feedback received was about the difficulty distinguishing between high and low marks within a grade boundary.

Semi-automated feedback tool

A concern with fully or semi-automated feedback is the loss of personalised feedback (Table 1). Most systems aim to help markers rather than replace their feedback, such as that of Buyrukoglu, Batmaz, and Lock (2019). Their system, used for marking programming code, allows markers to write and review comments. Earlier systems, like Hepplestone and Mather (2007), supplemented printed feedback with oral comments, requiring staff to remember or review feedback. Oral feedback may be preferred by some students in need of additional explanation (Zhang and Hyland, 2018). Leibold and Schwarz (2015) note the importance of body language in feedback, which is absent in electronic/online systems. One of the benefits of our tool is that it allows markers to add supplementary comments for personalisation (Figure 2).

Advantages	Disadvantages
Faster	Loss of personalisation
Reduced workload	Loss of detail
Cost saving due to less time needed for staff marking	Potential increase of workload if the system fails

Table 1. Advantages and disadvantages of adapting to an automated or semi-automated system

The screenshot displays a marking interface for a poster presentation. At the top, it says '1) Poster presentation..... / 25 marks'. Below this is a table with four columns. The first column is labeled 'The audibility during the presentation'. The second column contains a red box with the text 'Please put "x" in this column to reflect your feedback.' and a small red 'a' below it. The third column contains the text 'was clear.' and the number '5'. The fourth column is empty. To the right of the table is a large green box labeled 'Additional comments: Good however,.....'. To the right of the green box is a vertical bar labeled 'Total mark' with the number '0'. Below the table is a red box labeled 'Feedback' containing the text 'Good however,.....'.

Figure 2. This figure, shows a screenshot of a section from the semi-automated tool, also including the free-text section.

A concern with automation is the potential drop in feedback quality compared to manual marking. However, automated systems reduce marking time (Barker, 2011), allowing markers to review, edit and add comments. High-quality marking should maintain student satisfaction (Office for Students, 2021). Zhang and Hyland (2018) found that students were not used to receiving such detailed feedback, but staff time constraints make this impractical for all. Feedback could be adjusted to be more effective.

Nicoll, Douglas and Brinton (2022) explored effective feedback, noting that grade-related words positively impact student grades. Our system, like Barker (2011) and Hepplestone and Mather (2007), is based on the assignment marking scheme. This ensures feedback uses grade-oriented language, potentially improving student grades and helping them understand where they gained or lost marks. Additionally, Nicoll, Douglas and Brinton (2022) suggested that using students' names can improve performance, which could be trialled in other assessments. Despite marking policies requiring anonymity, our tool is planned for face-to-face poster marking where anonymity isn't feasible. Hepplestone and Mather (2007) proposed visual sliders to help students understand their position on the marking scale. Nicoll, Douglas and Brinton (2022) also suggested using names in feedback, similar to Barker's (2011) idea of incorporating student images. However, there's insufficient evidence on the benefits of visual sliders or images, apart from user-friendliness (Barker, 2011).

User-friendly marking tools are crucial in literature (Hepplestone and Mather, 2007). While some systems use separate staff/student interfaces (Batmaz, Stone and Hinde, 2019; Stanger, 2018), our tool doesn't need this due to the fact that feedback and grades are copied to the online learning platform. Other systems minimise copying through batch uploads (Hepplestone and Mather, 2007) or direct student interaction (Batmaz, Stone and Hinde, 2019; Stanger, 2018).

Semi-automated marking systems represent a promising approach to reducing workload (Batmaz, Stone and Hinde, 2019; Buyrukoglu, Batmaz and Lock, 2019; Hepplestone and Mather, 2007; Stanger, 2018). Similarity grouping of student feedback also allows for efficient application of comments across groups (Batmaz, Stone and Hinde, 2019;

Buyrukoglu, Batmaz and Lock, 2019). While these systems have demonstrated benefits in reducing repetitive marking tasks, particularly in fields with definitive “correct” answers such as programming, they still face limitations (Hepplestone and Mather, 2007; Stanger, 2018). The need for individual review of automatically generated feedback and manual intervention adds complexity to the process (Buyrukoglu, Batmaz and Lock, 2019). Further research and development is needed to enhance these systems' capabilities while maintaining assessment quality and reducing overall marking time.

References

- Anthonyamy, L, Koo, A C and Hew, S H (2020) 'Self-regulated learning strategies in higher education: Fostering digital literacy for sustainable lifelong learning', *Education and Information Technologies* 25: 2393-2414. <https://doi.org/10.1007/s10639-020-10201-8>
- Barker, T (2011) 'An Automated Individual Feedback and Marking System: An Empirical Study', *The Electronic Journal of e-Learning*. 9 (1): 1 – 14 Available at: <https://www.academic-publishing.org/index.php/ejel/article/view/1601> [accessed 20 July 2022].
- Batmaz, F, Stone, R and Hinde, C J (2009) *Personalised Feedback with Semi-automated assessment tool for conceptual database model*. Available at: https://figshare.com/articles/conference_contribution/Personalised_feedback_with_semi-automated_assessment_tool_for_conceptual_database_model/9403814 [accessed 20 July 2022].
- Brinton, C G and Chiang, M (2014) 'Social Learning Networks: A brief survey', in *2014 48th Annual Conference on Information Sciences and Systems (CISS)*, pp 1-6. <https://doi.org/10.1109/ciss.2014.6814139>
- Buyrukoglu, S, Batmaz, F and Lock, R (2019) 'Improving marking efficiency for longer programming solutions based on a semi-automated assessment approach', *Computer Application in Engineering Education*, 27 (3): 733-743.
- Calvo, R A and Ellis, R A (2010) 'Students' Conceptions of Tutor and Automated Feedback in Professional Writing', *Journal of Engineering Education*, 99 (4): 427-348.
- Dawson, P, Henderson, M, Mahoney, P, Phillips, M, Ryan, T, Boud, D and Molloy, E (2019) 'What makes for effective feedback: staff and student perspectives', *Assessment and Evaluation in Higher Education*, 44 (1): 25-36. <https://doi.org/10.1080/02602938.2018.1467877>
- Gonzalez, T, De La Rubia, M A, Hincz, K P, Comas-Lopez, M, Subirats, L, Fort, S et al (2020) 'Influence of COVID-19 confinement on students' performance in higher education', *PLoS One*, 15:e0239490. <https://doi.org/10.1371/journal.pone.0239490>
- Hepplestone, S and Mather, R (2007) 'Meeting Rising Student Expectations of Online Assignment Submission and Online Feedback', in Khandia, F (ed.) *11th CAA International*
-

Computer Assisted Assessment Conference: Proceedings, 10-11 July 2007, Loughborough University. Loughborough: Loughborough University.

INTO University Partnerships (2024) *Into survey shows strong satisfaction for UK programs*. Available at: <https://www.intoglobal.com/news/news-articles/2024/into-s-2024-annual-student-experience-survey-shows-strong-satisfaction-for-support-services-for-uk-programs/> [accessed 1 March 2025].

Leibold, N and Schwarz, L M (2015) 'The Art of Giving Online Feedback', *The Journal of Effective Teaching*, 15 (1): 34-46.

Nicoll, S, Douglas, K and Brinton, G (2022) 'Giving Feedback on Feedback: An Assessment of Grader Feedback Construction on Student Performance', in *LAK22, 12th International Learning Analytics and Knowledge Conference*, 21-25 March, United States of America. New York: Association for Computing Machinery.

Office for Students (2021) *National Student Survey 2021 results*. Available at: <https://www.officeforstudents.org.uk/data-and-analysis/national-student-survey-data/nss-data-archive/nss-2021-results/> [accessed 20 July 2022].

Office for Students (2022) *NSS Data: Provider-level*. Available at: <https://www.officeforstudents.org.uk/data-and-analysis/national-student-survey-data/nss-data-archive/nss-2022-results/> [accessed 20 July 2022].

Office for Students (2024) *NSS Data: Provider-level*. Available at: <https://www.officeforstudents.org.uk/data-and-analysis/national-student-survey-data/nss-data-archive/nss-2024-results/> [accessed 23 March 2025].

Royal Holloway (2024) *Top marks for student satisfaction in Geography*. Available at: <https://www.royalholloway.ac.uk/research-and-education/departments-and-schools/geography/news/top-marks-for-student-satisfaction-in-geography/> [accessed 1 March 2025].

Stanger, N (2018) 'Semi-automated assessment of SQL schemas via database unit testing', in Yang, J C, Chang, M, Wong, L-H and Rodrigo, M M T (eds) *International Conference on Computers in Education*, 26-30 November, Philippines. Taiwan: Asia-Pacific Society for Computers in Education.

Taghipour, K and Ng, H T (2016) 'A Neural Approach to Automated Essay Scoring', in Su, J, Duh, K and Carreras, X (eds) *Empirical Methods in Natural Language Processing*, 1-5 November, Texas. United States of America: Association for Computational Linguistics.

Zhang, Z and Hyland, K (2018) 'Student engagement with teacher and automated feedback on L2 writing', *Assessing Writing*, 36: 90-102.

Zhu, M, Liu, O L and Lee, H-S (2019) 'The effect of automated feedback on revision behavior and learning gains in formative assessment of scientific argument writing', *Computers and Education*, 143 (103668). Available at:

<https://www.sciencedirect.com/science/article/pii/S0360131519302210> [accessed 20 July 2022].

Appendix A

Citation	System used	How the system works	Year	Faculty specific
Calvo and Ellis	Glosser	Analyses text and then provides scored feedback	2010	Yes Engineering –
Batmaz, Stone and Hinde	Own	Groups similar student diagrams together to then be marked by the marker and then copied across the diagrams	2019	No
Buyrukoglu, Batmaz and Lock	Own	Puts segments of student programmes into groups and then marks them and uses the markers' comments for similar segments in that group	2019	Yes Programming –
Zhang and Hyland	Pigai	Analyses text and offers automated scoring	2018	Yes Language studies –
Zhu, Liu, and Lee	c-rather-ML	Automated scoring engine that provided scores on students written answers about climate change	2019	No
Hepplestone and Mather	Own – created in Excel	Matrix of statements which can be selected and used at the user's discretion	2007	No
Barker	Own	Based on a marking scheme created by the subject tutor	2011	No
Nicoll, Douglas, and Brinton	Own	Analysed feedback for textual information	2022	No but used for Engineering in study
Stanger	Oracle		2018	Yes Programming –

Creating authentic assessments for an AI-enabled world

Incorporating artificial intelligence ethically in authentic assessment design and assessment in higher education

Stephen Bunbury, Reader (Associate Professor) in Law, University of Westminster, UK

Stage one: background

Driven by technological advancements, higher education (HE) is undergoing rapid transformation. Unprecedented challenges, such as the pandemic, have highlighted the urgent need for interventions to explore innovative and adaptable approaches to student learning and assessment in HE. Artificial intelligence (AI) in education has become pivotal in reshaping assessment strategies. Moreover, the pandemic accelerated the use of different modes of learning. Since the shift to online and hybrid methods of delivering higher education during the pandemic, educators were forced to reevaluate traditional assessment approaches. The rapid integration of AI-assisted tools has further accelerated this transition, prompting a move toward more authentic, technology-enhanced assessment models that not only uphold academic integrity but also enhance accessibility and inclusive curriculum design. These shifts suggest a growing trend in the need to integrate AI responsibly into educational assessment frameworks, ensuring alignment with principles of equity, ethics and engagement.

Advancements in technology and competency-based education have supported the adoption of AI (Darling-Hammond and Adamson, 2014). Nevertheless, barriers remain, including resource demands, concerns about reliability, and lack of staff training (Gulikers et al, 2004). Over time, authentic assessment has developed from viewing authenticity in assessment as replicating real-world tasks to fostering societal values such as social justice, inclusivity and wellbeing (McArthur, 2023; Ajjawi et al, 2023). The rise in AI assistance has enabled flexible and accessible assessments. However, this challenges compatibility with traditional learning environments (Vlachopoulos and Makri, 2024).

To address this, our institution seeks to implement, replicate and adapt methodologies inspired by Bristol University's exemplary practices in assessment design across disciplines. These practices are particularly relevant for diverse student populations, such as at the University of Westminster, which includes students from various nationalities and a broad range of socio-economic and cultural backgrounds.

Bristol University's Philosophy module, which has successfully implemented authentic assessment techniques, such as simulation-based tasks and peer evaluations, has served as a pivotal model for innovation. Drawing inspiration from the Philosopher Queens unit, the module challenges traditional, Eurocentric and male-dominated perspectives in philosophy. By emphasising the systemic exclusion of women in the discipline, as highlighted by Dietzel (2024), the course enriches the learning experience, addressing critical issues of inclusivity and representation in philosophical studies. Encouraging students to critically engage with AI's inherent biases (particularly those reflecting societal prejudices about race and gender)

provides a framework for understanding how technology perpetuates these imbalances (Dietzel, 2024). The emphasis on reflection and critical use of AI aligns with Bristol's commitment to fostering deeper, context-driven learning experiences. Although Bristol is a Russell Group university, the insights gained from this pilot are transferable to any HE institution seeking to integrate AI ethically. The Bristol pilot, conducted with the support of the study skills team, featured sessions co-created and delivered in collaboration with academic staff and students. It allowed students to practice using AI to explore societal biases. As part of their learning journey, students demonstrated their understanding by creating images illustrating specific philosophical biases.

Philosophers Queens examples of images (Dietzel, 2024)

Building on these principles, implementing inclusive, practical, real-world assessment methods that align with the needs of today's learners remains a key priority at the University of Westminster. Historically, assessment in HE has relied heavily on traditional methods

such as rote memorisation and knowledge reproduction. These approaches often fail to cultivate critical thinking and creativity, skills essential for real-world application. As defined by Gulikers et al (2004), authentic application emphasises realism, cognitive challenge and evaluative judgment. Such methodologies are increasing as the rise of generative AI tools such as ChatGPT necessitate assessments that encourage critical engagement and originality.

To address concerns about academic integrity in this AI-driven era, the initiative focuses on creating unique, context-specific, authentic assessments less susceptible to manipulation by generative tools. Bristol's Philosophy module demonstrates best practices by incorporating simulation-based tasks, peer evaluations and oral components. Inspired by this pilot's success, our institution intends to pilot similar approaches across disciplines other than law.

Stage two: approach

Integrating authentic assessment at the University of Westminster was inspired by Bristol University's Philosophy module for authenticity. The critical element of this module emphasises designing assessment tasks that reflect real-world challenges, embedding these tasks within meaningful and relevant contexts, fostering collaboration to simulate professional social environments, ensuring that assessment results align with industry and professional standards, and grounding evaluation criteria in practical and applicable terms.

This approach seeks to enhance assessment relevance and impact by aligning tasks with real-world challenges and professional standards. While the core framework (designed to make assessments more relevant and impactful by closely mirroring real-world scenarios and professional expectations) remains consistent, its implementation varies across disciplines.

For example, in applied sciences and engineering, assessments often involve problem-based case studies that simulate industry challenges. In contrast, law requires a different set of skills and a distinct approach, where assessments frequently focus on the application of law, statutory interpretation and problem-solving through structured legal arguments. While traditional law assessments emphasise skills developed through written assignments, problem questions and mooting, there is a growing need to embed more authentic, practice-based assessments in curriculum design. For instance, simulated client interviews, legal clinic work, or negotiation exercises could provide real-world applications of legal skills while still aligning with the discipline's learning outcomes. Ensuring that assessments remain discipline specific while allowing for the adaptability of the framework will help create assessments that reflect real-world practice in professional settings across various disciplines. Our pilot incorporates reflective tasks that focus on critiquing AI and reflecting on a task produced by AI, taking a reflective approach that enables students to apply theoretical knowledge to practical scenarios that include an element of reflection.

To promote AI-driven authentic assessment practices, a series of dissemination activities were undertaken across 2024 and 2025, beginning with a presentation in September 2024 at the University of Westminster Learning and Teaching Symposium. This session, titled

Assessment Design and the Use of Artificial Intelligence: How Can Assessments Be Meaningful Whilst Maintaining Academic Integrity?, explored the role of AI in shaping assessment approaches while ensuring academic integrity.

Building on this, in November 2024, the approach was further disseminated at Advance HE's Assessment and Feedback Symposium 2024 in a presentation titled Incorporating Artificial Intelligence Ethically in Authentic Assessment. This session emphasised the ethical integration of AI tools into assessment design to support authentic learning experiences.

Following these engagements, I was invited to further share this work in November 2024 through dissemination in the PGCert in HE programme, fostering discussions among early-career academics. Recognising its growing impact, I was subsequently invited back in March 2025 to continue discussions and further develop its integration.

These efforts culminated in 2025 with an invitation to present at a university-wide Lightning Talk as part of the Centre for Education and Teaching Innovation. This session, titled Designing Meaningful Assessments in the Age of Artificial Intelligence: Balancing Innovation with Academic Integrity, aimed to expand institutional dialogue on AI-driven assessment design and ethical considerations.

As seen in the diagram, authentic assessment addresses challenges such as academic integrity while emphasising critical thinking and collaboration. A symposium on this topic provided a platform for staff engagement, leading to an invitation to contribute to the Postgraduate Certificate in Higher Education (PGCertHE) programme. Within this framework, sessions on authentic assessment design focus on key principles such as realism, student collaboration and ethical AI use. These sessions aim to equip educators with practical strategies for designing assessments that enhance originality and deeper

learning while addressing challenges posed by generative AI tools. During my tenure as an Academic Professional Development Fellow at the University of Westminster's Centre for Education and Teaching Innovation and through contributions to the PGCert programme, significant efforts were initiated to embed authentic assessment into professional development at the University of Westminster. For example, workshops with early-career academics redesigned learning outcomes, integrating critical AI evaluation, rubrics, and student reflections on ethical and responsible AI use.

Stage three: outcomes

These initiatives produced impactful outcomes. Staff workshops on academic integrity and authentic assessment design provided practical guidance on creating assessments resilient to AI misuse. These workshops effectively mitigated the risks of academic dishonesty by focusing on tasks requiring a unique, real-world application. Additionally, contributions to the PGCert in HE established a sustainable model for embedding authentic assessment principles into staff training. These sessions promoted a reflective and collaborative approach to designing innovative assessments.

Broader dissemination across the university was achieved through interactive university-wide CPD workshops that encouraged staff to engage with and apply authentic assessment principles tailored to their specific disciplines and needs. An example of dissemination at the University of Westminster Learning and Teaching Symposium 2024 involved an interactive workshop to engage staff with authentic assessment principles and explore their application across various disciplines. One notable exercise invited participants to use AI to generate a 500-word essay on 'Should cannabis be legalised?' They then summarised the content in their own words within a 100-word limit and composed a 500-word reflective response on the key skills developed through the activity. This exercise highlighted the flexibility and relevance of authentic assessment principles, demonstrating their applicability across diverse academic contexts beyond traditional disciplines such as philosophy or law.

Integrating such practices into staff development programmes established a foundation for transformative assessment design at the university. These workshops address contemporary challenges such as academic integrity and equip students with critical skills for navigating an AI-influenced professional landscape. The reflective exercise highlighted how authentic assessment promotes creativity, adaptability and skill-building, contributing to a long-term, sustainable model for innovative education.

References

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.

<https://doi.org/10.1080/02602938.2023.2271193>

Darling-Hammond, L and Adamson, F (2014) *Beyond the bubble test: How performance assessments support 21st century learning*. Bognor Regis: John Wiley and Sons.

Dietzel, A (2024) 'Using AI as a teaching tool for exploring societal biases'. Bristol University PowerPoint slide presentation. Bristol Institute for Learning and Teaching Conference, held on 10 July 2024. U

Gulikers, J, Bastiaens, T J and Kirschner, P A (2004) 'A five-dimensional framework for authentic assessment', *Educational technology research and development*, 52 (3): 67-86. <https://doi.org/10.1007/BF02504676>

McArthur, J (2023) 'Rethinking authentic assessment: work, well-being, and society', *Higher Education*, 85: 85-101. <https://doi.org/10.1007/s10734-022-00822-y>

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.

Authentic assessment design in science – a case study of three approaches

Rosanne English, Principal Teaching Fellow; Ainsley Miller, Teaching Fellow; Felicity Carlysle-Davies, Kristen Ceniccola-Campos, Knowledge and Exchange Associate, Strathclyde University, UK

Stage one: background

Assessment is at the core of higher education: it is a key source of motivation for learning (Boud, 1995) and ensures students have achieved the intended learning outcomes. However, assessment presents challenges as educators aim to ensure assessment is fair, valid, reliable, inclusive and authentic. Authentic assessment aims to ensure that assessment methods are valid and reflect real world tasks such that students are prepared for their careers post education. However, there are several challenges in developing such assessments including resistance to updating assessment practices and generative AI (genAI). Generative AI has resulted in some retreating to traditional exams, which are criticised for value (Lawerys, 1969) and reliability (Ramsden, 1992). In this case study, we present three examples of genAI resilient authentic assessment within science subjects at the University of Strathclyde in undergraduate and MSc modules in computer science, forensic chemistry and mathematics and statistics.

Stage two: approach

The following subsections provide an overview of the approaches of scenario-based assessment in computer science, continuous assessment in mathematics and statistics, and video and calculation-based assessment in forensic science.

Scenario-based assessment in computer science

Given the aim of balancing authentic assessment with mitigation of use of genAI tools, one approach is to adapt assessment to rely less on the final product and more on emulating real-world scenarios and assessing the process and higher order thinking skills which are not so easily replicated using genAI tools.

To achieve this, in computer security modules we have implemented these principles by incorporating recent real world case studies. Using recent real-world cyber security incidents supports students in exploring an authentic scenario in which there is often some information available but also requires critical thinking to identify what possibly led to a given cyber security failure, and what would assist in avoiding such a problem in future.

This helps students synthesise the information from the module and consider how these concepts integrate into real life. The assessment itself requires students to research and reflect on the incidents and write an executive summary which generally reflects on what went wrong and how could it be fixed.

It's important to note that many of these incidents do not have fully documented details available in online sources; they require critical thought to be applied to identify where vulnerabilities may have crept into a system, collating information from a range of resources, and presenting a reflective argument on what would be an appropriate way to mitigate such issues.

To implement this approach, we first identify appropriate real-world cases which can be used. A review of recent incidents is then undertaken to identify examples of relevant attacks. A short list of possible cases is identified, and scoping is undertaken to ensure suitability. A key component of engaging students is to provide them with the opportunity to select their case, which can be facilitated through the virtual learning environment.

As well as incorporating real world cases, the assessment rubric must also align with the expected focus on analysis and critical thinking. This is achieved by explicitly incorporating these as criteria, such as including 'evidence of research' showing the culmination and synthesis of a range of sources to support arguments in the given context. To support students in understanding what is expected, support can be provided through evaluation of past submissions against the marking criteria.

Continuous assessment in mathematics and statistics

There has been limited innovation in the teaching of mathematics, with many lecturers continuing to rely on a "chalk and talk" approach. To address challenges posed by genAI tools in mathematics, our institution has reinstated traditional closed book written exams. While genAI struggles with more theoretical aspects of mathematics, students who overly rely on such tools may lack a deep understanding of fundamental concepts, ultimately hindering their progress in subsequent study.

Initially, a shift back to closed-book was met with worry from our undergraduate students, many of whom had become accustomed to remote open book exams during the pandemic. To ease the transition for students, formulae sheets were introduced and were made available at least three weeks before the examination. The introduction of formulae sheets has been widely praised by students, with students anecdotally saying that the examinations feel less like a "memory test" and are more like a genuine assessment of their mathematical ability. This focus on reinforcing mathematical theory lays the groundwork for future assessments that might involve critiquing AI outputs.

In contrast, statistics modules incorporate continuous assessments. Final examinations typically take place in computer laboratories, where students are also provided with a coding reference sheet. Additionally, in some modules, students are permitted to access the internet during the examination to troubleshoot coding errors, as this is more authentic. However, genAI is prohibited in final and continuous assessments. This consistent approach to statistics assessments has enabled students to develop practical coding skills and adapt to real-world challenges without becoming overly dependent on AI.

Implementing the formulae and/or coding sheets for both mathematics and statistics examinations is not resource intensive. Typically, these sheets are created mid-way through

the semester and uploaded to the virtual learning environment. Students are encouraged to use them during computer laboratory sessions to increase familiarity. Some lecturers also discuss potential gaps in the coding sheet with students or explain the rationale for omitting certain key code or formulae.

Implementing continuous assessment requires more careful planning. One of the main challenges is sourcing a suitable dataset or context that reflects a real-life scenario, ensuring authenticity and relevance. Continuous assessment can also increase the marking burden for staff, although automated marking can help. Care must also be taken to ensure that the same skill is not assessed in both the project and the examination, which could result in students being penalised twice for the same gap in knowledge.

Video-based and calculation-based assessments in forensic chemistry

Forensic science is a vocational subject with a real need for authentic assessments to support student learning. Within the forensic chemistry specialisation two areas were identified that could relatively easily be adapted from exam hall-based assessments to continuous assessments focusing on forensic drug analysis and toxicological calculations.

In an effort to offer a practical, career-focused assessment scenario, students were provided with a written prompt and asked to critically evaluate and discuss a plausible and hypothetical outcome for the forensic drug analysis topic. While their intended careers involve significant time spent in the laboratory, a critical aspect of student learning should focus on strengthening their delivery of evidential findings to a wider audience, including fellow scientists, law enforcement and legal personnel involved throughout their examination process. Students were tasked with creating three-minute videos specifically targeting each of the noted audiences using terminology and descriptions appropriate for each. Although students expressed an initial nervousness at the challenge of recording themselves talking in front of a camera, their feedback suggested they enjoyed a change from traditional examinations and appreciated an early opportunity to prepare for real-world experiences.

To implement this approach, consideration of current analytical challenges is needed. The initial step is to identify a suitable type of forensic drug evidence that will require a multi-step analytical procedure. Secondly, analogies and terminology must be identified that is appropriate for each step of the procedure. Students are provided with an opportunity to practice delivering these analogies and scientific terms, which improves their confidence in describing various scenarios prior to being presented with the assessment prompt.

For the calculations-based assessment, the calculations associated with blood and breath alcohol level determination and drug elimination were transferred to an online open book timed assessment. This allowed for a wider range of questions to be asked and more closely represents a real-world situation where a forensic scientist might be asked to review the data but would not be required to have the equations memorised. The assessment structure was co-created with the students, who preferred the option to upload handwritten answers and have more shorter questions rather than fewer long calculations. A quiz format was used in the virtual learning environment where a question bank was created to ensure

that students would not receive the same set of questions to help reduce collusion. While this was a significant initial outlay of staff time, it means that the questions can be reused each year as the questions for each student are randomised. The quiz was set so that students could access it at any time during a 12-hour time period but once they started the quiz they had a fixed time to complete the assessment. The students performed well in the assessment, but similar mistakes were seen when compared to an exam, suggesting that moving to an open book assessment has not made the assessment easier and continues to challenge students to apply their knowledge to the calculation-based problems.

Stage three: outcomes

While it can be daunting to adapt and change a traditional approach to assessment, students across all three interventions have responded well to these approaches, which aid in diversifying assessment, an approach known to improve outcomes for learners (French, 2024).

We believe the approaches presented are transferable to different contexts. Real world contexts and key skills and knowledge can be adapted to different disciplines. Recorded presentations offer the opportunity for students to evidence communication skills as appropriate; in graduate jobs graduates will likely be asked to speak about their work. As such, this is an opportunity for students to develop these skills in a less pressured environment by recording themselves, which also offers a chance to reflect more on their presentation with a recording being available when compared to a live presentation. Video interviews are also becoming an increasingly common part of recruitment processes across disciplines, providing students with an opportunity to practice effective communication in this format.

The key messages we believe can be taken from these case studies are:

- + Addition of real-world contexts both reduces the capacity for misuse of generative AI and provides a wider perspective of the skills being developed.
- + Key skills can be incorporated into continuous assessment to ensure students have sufficient hands-on experience in a low-stakes environment, improving assessment literacy.
- + Assessment techniques, which are not often discussed in a science context, can be adopted.

Currently, evidence is anecdotal. We plan on formally evaluating the student experience of these approaches when compared with traditional assessments. Furthermore, we hope to evaluate the long-term impact on student experience and outcomes. We would also be interested in hearing from colleagues should they chose to adapt similar approaches, and welcome discussion.

References

- Biggs, J (2001) 'The reflective institution: Assuring and enhancing the quality of teaching and learning', *Higher Education*, 41 (3): 221-238. <https://doi.org/10.1023/A:1004181331049>
- Boud, D (1995) 'Assessment and Learning: Contradictory or Complementary?', in Knight, P T (ed.) *Assessment for Learning in Higher Education*. London: Kogan Page, pp 35-48.
- French, S, Dickerson, A and Mulder, R A (2024) 'A review of the benefits and drawbacks of high-stakes final examinations in higher education', *Higher Education*, 88: 893-918. <https://doi.org/10.1007/s10734-023-01148-z>
- Lauwerys, J A and Scanlon, D G (eds) (1969) *World Yearbook of Education 1969: Examinations*. 1st edn. Abingdon: Routledge. <https://doi.org/10.4324/9780203080597>
- Ramsden, P (1992) *Learning to teach in higher education*. Abingdon: Routledge.
- Villarroel, V, Bloxham, S, Bruna, D, Bruna, C and Herrera-Seda, C (2017) 'Authentic assessment: creating a blueprint for course design', *Assessment and Evaluation in Higher Education*, 43 (5): 840-854. <https://doi.org/10.1080/02602938.2017.1412396>

Authentic assessment: the current and future role of AI in marketing strategies

Lucill Curtis, Associate Professor in Marketing, Norwich Business School, University of East Anglia, UK

Introduction

This is a case study focusing on the integration of generative AI (GAI) into authentic formative and summative assessments, as part of an MBA module, Marketing and Customer Focus. The aim is to ascertain, develop and align candidate's existing GAI capabilities with the requirements of 'real world' companies developing marketing strategies. The students on the module are a combination of part-time executive MBA and senior level apprenticeship (SLA) candidates.

Stage one: background: integration of AI into marketing strategies

A particularly significant disruptor in recent years for marketing and marketers has been the persistent integration of new technology and applications, especially those enabled by artificial intelligence (AI) (Grewal, Motyka and Levy, 2018). Artificial Intelligence Marketing is an approach designed to maximise the use of technology and market data in order to improve the customer experience (Jain and Aggarwal, 2020). In particular, AI is understood as a vital contributor to marketing efficiency and brand identity, according to an online survey of 1,200 marketers conducted by Wakefield Research (AdAge, 2024). AI in marketing encompasses a broad range of tools and technologies designed to enhance various aspects of marketing strategies (Huang and Rust, 2021). In particular, GAI have started to play a significant role in producing creative elements for marketing, such as writing copy, designing visuals, and even generating music or video content. GAI is one type of AI that can be used to create new content based on large volumes of data from a variety of sources. ChatGPT, Microsoft Copilot and Google Gemini are GAI tools, built on large language models (LLMs) (Department for Education, 2025).

These tools allow marketers to produce large volumes of content quickly, ensuring that they can respond to market trends and customer demands with unprecedented speed and efficiency (Getchell et al, 2022). ChatGPT can also make (substantial) errors and exhibit bias. A recent policy paper from the Department for Education outlines the vast potential of AI in education to encourage innovation. The UK government is working with the education sector, educational technology (edtech) industry, experts and academics to build evidence and support the education sector to use AI safely, responsibly and effectively. From their initial research and engagement with the sector, they suggest GAI could be used for: creating educational resources, lesson and curriculum planning, tailored feedback and revision activities, administrative tasks and supporting personalised learning (Department for Education, 2025).

Arguably GAI is one of the more straightforward aspects of the AI portfolio to use. For marketers, the impact of ChatGPT's success has gone beyond the norm, bringing about more profound opportunity and realisation, making work easier, more interesting and fun (AdAge, 2024). However, despite positive improvements presented by AI, nearly 9 in 10 marketers expressed concern that *"I will have to learn how to use AI to maintain my value as a marketer"* (AdAge, 2024, 3). Thus, Guha, Grewal and Atlas (2024) suggest that, by marketing educators including ChatGPT in their lessons, they can enhance learning experiences and better prepare students for future jobs with marketing firms that rely on GAI in practice.

Within current scholarship there is an emphasis on overcoming the negative uses of GAI in assessment writing by students, which threatens the veracity of student work and therefore also the authenticity of assessments (Hack and Knight, 2023). For example, a recent systematic literature review of trends, adoption measures, diverse applications and current limitations of GAI (ChatGPT) research in higher education identified academic integrity, ethics and student productivity as the main themes in 2023-24 (Baig and Yadegaridehkordi, 2024). Therefore, with the emphasis primarily on the less positive implications of GAI, this case study draws specifically on authentic learning pedagogy to design formative and summative assessments to promote beneficial adoption and learning (Herrington, Reeves and Oliver, 2010). Authenticity has a multitude of meanings in learning and curriculum scholarship. In this context it can be viewed as aligning students' assessment tasks and learning with real-world practice (Stein, Isaacs and Andrews, 2004). Therefore the aims and objectives were:

- + to ascertain MBA/SLA students' current experience and application of GAI, including ChatGPT
- + to empower and build confidence for MBA students to use GAI to design marketing communications strategies and plans in order to test the efficiency principle for marketers
- + to identify issues with GAI in marketing practice, relating to ethics, reliability, personalisation and empathetic messaging
- + to make recommendations for future GAI enhancements/improvements to enhance take up and implementation in marketing.

Specifically, by including GAI in their lessons, marketing educators can enhance learning experiences and better prepare students for future jobs with marketing firms that rely on ChatGPT in practice at whatever career stage they are at (Guha, Grewal and Atlas, 2024).

Stage two: approach to authentic assessment

The Marketing and Customer Focus module has four six-hour workshops, plus six 1.5 hour online sessions via the web-based video conferencing tool Blackboard Collaborate. This case study focuses on the authentic assessments embedded during teaching in March and April 2024. The assessment was aligned to the Gulikers, Bastiaens and Kirschner (2004)

five-dimensional framework for authentic assessment (see Figure 1, below). Their definition of authentic assessment sets out a concept that requires students to use the same capabilities or blend of learning, abilities and attributes they need to carry out the role professionally in an industry (marketing) context.

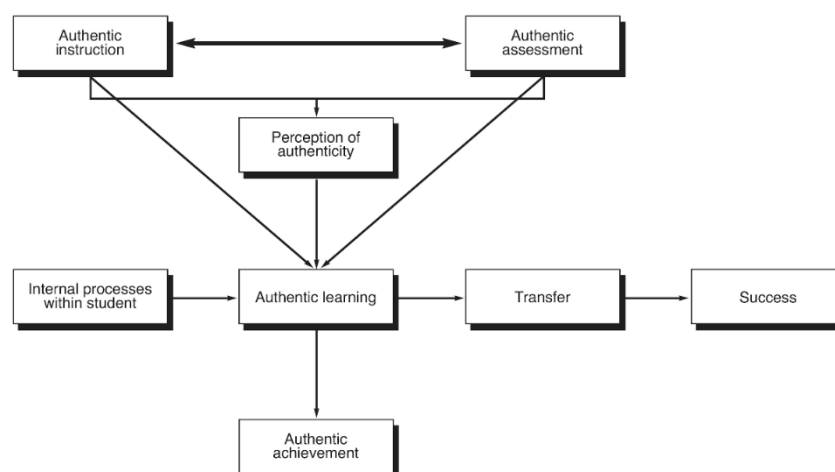


Figure 1: General framework for authentic assessment (Gulikers, Bastiaens and Kirschner, 2004, 70).

This module, which runs every year in semester two (March-April), has been designed to cover a broad introduction to marketing for professionals from a range of sectors, including health, local government and the private sector. The majority of the 35 students have limited experience and expertise related to marketing. The module content therefore outlined the basic theoretical concepts, such as the marketing mix, segmentation, targeting and positioning (STP) as well as the impact and integration of various technologies, including GAI, into marketing practice. For example, lecture content highlighted the efficient way insights can be generated by GAI, through the personalising of content and offerings. The decision to incorporate GAI components was a natural part of annually refreshing course content, in line with developments in marketing practice. Current research claims that marketing content generated by GAI is likely to be more individually relevant than that produced by previous digital technologies (Kshetri, Dwivedi, Davenport and Panteli, 2024). Assessments for this module comprised two group formative (practice) tasks, which all students completed in class, and one individual summative task, which were validated internally before the module commenced. Thus, the rationale for the authentic assessment strategy was to enable students to experiment with GAI tools in groups through the formative tasks, in a way that reflects marketing practice, alongside the integration of marketing theories into possible real world scenarios – whereas the summative assessment was an individual authentic piece of work based on the students designing a marketing strategy and plan to encourage growth for their own business/organisation. Students were encouraged within the summative assessment to suggest ways where the integration of GAI could make the process of writing a marketing strategy and plan more efficient and

impactful; for example it is a quick and useful way of identifying key headings for a marketing strategy.

The formative assessments were undertaken during the first two face-to-face sessions. The approach focused on teamwork, collaboration and knowledge sharing through group presentations. Discussion and questions followed from peers and the academic module organiser. Two different marketing practitioners also helped to write the formative assessments, which were on different days. The practitioners' experience and evaluations were incorporated into the formative assessment feedback. Professional activities incorporated within the assessment exposed these students to practice-based real-world activities and challenges, thereby supporting their development and better equipping them for the summative assessment and relevant industry roles (Vohmann, 2019).

Formative one: students worked in groups of four or five and developed an organisation from the travel sector, which could be based on an existing company. They were asked to list five key segments, then to choose a segment most likely to produce growth in that calendar year, as well as a targeting/positioning strategy. They were encouraged to use GAI as a research tool and a tactic to be integrated into the marketing planning. It was then their decision how to position the organisation to the chosen segment, including key messages, a brand identity, logo, values and messaging. Groups fed back to the class as a video blog, presentation or visual.

Formative two: within groups of four or five, students developed a marketing communications' plan for a new product, such as a health food bar, sportswear clothing range or restaurant. They were also allowed to focus on an area of their own companies. The first stage included using GAI as a research tool to identify a target segment, produce a persona and detail the key elements of a marketing campaign, ie objectives, key messages, channels, imagery and branding. Students produced one mocked up advert and storyboard for a video (or short video). They were encouraged to experiment with GAI as part of their preparation and final presentation. They were also asked to acknowledge and discuss its inclusion. The module organiser, peers and the practitioners – a brand manager and a creative director – provided verbal feedback immediately after the presentations. This was followed up with written feedback for each group, based on a marking rubric covering the core components of the assignment. Feedforward comments were also included to help candidates plan for the individual summative assessment.

Stage three: outcomes

This combination of authentic assessments within the level seven module Marketing and Customer Focus worked in encouraging students to share their experience and knowledge of applying GAI in their current roles, which helped them to share prior learning and build confidence. Students were then able to apply this knowledge to the marketing context authentically through their identification of potential issues and recommendations for case-study examples in the formative assessments and within their own company context, as part of summative requirements.

There was strong evidence of authentic/real-world learning having taken place. For example, more than a third (13 out of 35) of the students highlighted the effectiveness of using GAI as a research tool, as well as providing the structure and possible headings for marketing communications' strategies and plans. Those with limited marketing experience found this aspect particularly helpful and insightful. It also helped to overcome some earlier concerns about ethical issues relating to academic plagiarism, in particular, as headings were identified by GAI, but the content was researched and written by the students. The timesaving effectiveness of GAI was thus authentically proven.

Clear links were made to the relevance of these skills to current marketing practice, thus closing the authenticity loop, with one student stating that sessions and assessments "*were well planned, interesting, well delivered, and relevant.*" In order to broaden the integration of GAI across the teaching portfolio within Norwich Business School (NBS), starting with the rest of the MBA modules, I have been sharing my experience with colleagues at modular, course and programme level, through relevant team meetings. Many other colleagues are also starting to experiment with AI in teaching content, assessments and as part of support materials on the virtual learning platform (Blackboard). The head of school has also set up an AI taskforce to review best practice AI integration across undergraduate and postgraduate teaching. I have been able to contribute to the taskforce and ongoing innovation by sharing the planning and delivery of this module. I am intending to develop the assessment within this module further next year by providing more real world case-study examples demonstrating successful GAI integration into marketing practice, so these examples can also be built on as part of formative and summative assessments.

The next stage is to build on the integration of GAI into other projects and initiatives both within the University and externally within the wider community. This will be achieved through a technology focused pedagogy workshop within the business school, open to the other seven schools across the social sciences faculty at UEA. This will be an opportunity to share knowledge and experience more widely, regarding how academics are integrating technology, including GAI, into teaching, assessments and scholarship. Low levels of digital literacy (Ueno, Dennis and Dafoulas, 2023), are also prevalent in the local community and therefore academics within NBS are working with a local charity and practitioners to run community-based workshops to improve GAI literacy levels. These initiatives will help to broaden awareness, accessibility and implementation of GAI in more scenarios.

References

AdAge (2024) *Less do, more think: How to succeed in the AI-powered marketing era*. Available at: [2024-0131 Iterable-AdAge AI Whitepaper.pdf](#) [accessed 7 March 2025].

Baig, M I and Yadegaridehkordi, E (2024) 'ChatGPT in the higher education: A systematic literature review and research challenges', *International Journal of Educational Research*, 127, p 102411.

Department for Education (2025) *Generative artificial intelligence (AI) in education*. Available at: <https://www.gov.uk/government/publications/generative-artificial-intelligence-in-education/generative-artificial-intelligence-ai-in-education> [accessed 7 March 2025].

Getchell, K M, Carradini, S, Cardon, P W, Fleischmann, C, Ma, H, Aritz, J and Stapp, J (2022) 'Artificial intelligence in business communication: The changing landscape of research and teaching', *Business and Professional Communication Quarterly*, 85 (1): 7-33.

Grewal, D, Motyka, S and Levy, M (2018) 'The evolution and future of retailing and retailing education', *Journal of Marketing Education*, 40 (1): 85-93.

Guha, A, Grewal, D and Atlas, S (2024) 'Generative AI and marketing education: What the future holds', *Journal of Marketing Education*, 46 (1): 6-17.

Gulikers, J, Bastiaens, T and Kirschner, P (2004) 'The Five-Dimensional Framework for Authentic Assessment', *Educational Technology Research and Development*, 52: 67-86.

Hack, K and Knight, C (2023) *Higher education in the era of AI*. Available at: <https://www.advance-he.ac.uk/news-and-views/higher-education-era-ai> [accessed 7 March 2025].

Herrington, J, Reeves, T C and Oliver, R (2010) *A Guide to Authentic E-Learning*. New York: Routledge.

Huang, M H and Rust, R T (2021) 'A strategic framework for artificial intelligence in marketing', *Journal of the Academy of Marketing Science*, 49: 30-50.

Jain, P and Aggarwal, K (2020) 'Transforming marketing with artificial intelligence', *International Research Journal of Engineering and Technology*, 7 (7): 3964-3976.

Kshetri, N, Dwivedi, Y K, Davenport, T H and Panteli, N (2024) 'Generative artificial intelligence in marketing: Applications, opportunities, challenges, and research agenda', *International Journal of Information Management*, 75, p 102716.

Stein, S J, Isaacs, G and Andrews, T (2004) 'Incorporating authentic learning experiences within a university course', *Studies in Higher Education*, 29 (2): 239-258.

Ueno, A, Dennis, C and Dafoulas, G A (2023) 'Digital exclusion and relative digital deprivation: Exploring factors and moderators of internet non-use in the UK', *Technological Forecasting and Social Change*, 197, p 122935.

Vohmann, B L (2019) *Authentic assessment to enhance undergraduates' learning and development as effective practitioners in built environment disciplines*. PhD thesis. Anglia Ruskin University. Available at: [https://aru.figshare.com/articles/thesis/Authentic assessment to enhance undergraduates learning and development as effective practitioners in built environment disciplines/23763570?file=42199131](https://aru.figshare.com/articles/thesis/Authentic%20assessment%20to%20enhance%20undergraduates%20learning%20and%20development%20as%20effective%20practitioners%20in%20built%20environment%20disciplines/23763570?file=42199131) [accessed 7 March 2025].

Maintaining academic integrity in assessment and feedback

A multi-pronged and gamified approach to plagiarism prevention and academic integrity training

Laura Sharp, Professor of Pedagogy; Eric Davies, Lecturer; Mia Wilson, Lecturer; Ailsa Foley, Lecturer; Julie Langan Martin, Professor of Psychiatry; Dimitar Karadzhov, Lecturer, University of Glasgow, UK

Stage one: background

Like many institutions, the University of Glasgow has observed an increase in plagiarism cases since 2020 (Loveridge et al, 2024). Contributing factors likely include the impact of the pandemic, the shift towards online learning environments, and the emergence and rapid evolution of generative artificial intelligence tools (GAIT) (Ives and Cazan, 2024).

Within our student body, some concerning patterns supported by internal academic misconduct referral data include (Loveridge et al, 2024):

- + postgraduate students are more likely to be referred for plagiarism than undergraduates
- + mature students are more likely to be referred for plagiarism compared to younger peers
- + EU and other international students face a higher risk of plagiarism compared to their UK counterparts.

Such data underscores the need to design plagiarism prevention training targeted towards the distinct and situational needs and experiences of different student sub-cohorts (Tran et al, 2022). For example, cultural differences have been reported in beliefs about what constitutes academic dishonesty, including copying sentences and collaboration on coursework (Fatemi and Saito, 2020).

Stage two: approach

To support diverse student needs, our team of academics, alumni and a learning technologist designed an interactive, evidence-based online module – Academic Values, Originality, and Plagiarism – using [Articulate Rise](#). The module is innovative in its multi-pronged, eclectic approach. It presents more than 20 learning steps across four sections targeting academic integrity and values, understanding plagiarism, paraphrasing skills, and common pitfalls and good practice. The content includes university rules and consequences, expert videos, student testimonials, quizzes and game-based learning. It can be completed in under an hour and is delivered asynchronously for self-directed learning.

Instead of focusing solely on deterrence and compliance, the module addresses the root causes of plagiarism and contributing factors, such as misconceptions and skill gaps, thereby directing students to avoid unhelpful habits. Pedagogical research informed the inclusion of content targeting 'pain points', such as culturally mediated text borrowing practices, a lack of confidence in writing (authorial) identity, and inadequate engagement in

deep learning (Amsberry, 2009; Elander et al, 2010; Fatemi and Saito, 2020; Hayes and Introna, 2010; Tindall et al, 2021). Table 1 offers examples of included guidance.

Pain point	Example guidance
Culturally mediated text borrowing beliefs and practices	In some cultures, copying authors' words is considered respectful, while criticism is perceived as disrespectful (Hayes and Introna, 2010). Although direct quotations can sometimes be beneficial, they must be appropriately referenced, and critical analysis and original contributions must be demonstrated.
Lack of confidence and poorly developed writing identity	Students with low self-confidence as authors are at greater risk of plagiarism. Taking ownership of writing can minimise that risk. Students are asked to reflect on questions including: <ul style="list-style-type: none"> • How strongly do you see yourself as an author in relation to your university assignments? • What influences whether you develop a sense of authorship and ownership of your work?
Inadequate engagement in deep learning	Students are encouraged to focus on understanding the material rather than reproducing lecturers' or authors' words. Deep learning involves actively engaging with the material, connecting new learning with existing knowledge, and thinking critically about the information (Leenknecht et al, 2019).

Table 1. Guidance focused on text borrowing, writing identity and deep learning

Next, we outline the most innovative component of the module – a game-based learning resource, which opens the module.

Gamification

Games trigger a positive feedback loop upon task completion, promoting engagement (Kaufmann, 2018), and gamified audio-visual content can encourage active learning (Khan et al, 2017). As such, a 'Choose your own adventure'-style game was built using GAIT and incorporated into the module in October 2023 (Access game: [Academic Integrity and Values, Step 1](#)). Completion time is 10 min.

Developed over four weeks, the game presents two students in a coffee shop discussing six academic challenges, which reflect module discussion points about, for example, effective paraphrasing, responsible student collaboration when preparing assignments, ethical GAIT

use, and best practice when extracting details from papers (see Figure 1). Learners are prompted to make decisions on behalf of the characters about how to address the challenges. Each choice leads to distinct outcomes, demonstrating the immediate and long-term consequences of actions related to academic honesty, citation practices and research methodologies. The game supports learners to internalise values by observing outcomes in a virtual, consequence-free setting.

This innovative and budget-friendly resource introduced a steep learning curve, and feedback was sought from students, alumni, and teaching and digital colleagues throughout the design process to optimise its usability and value. Table 2 highlights various unexpected challenges faced during development. These issues underscored the complexities of using GAIT to authentically represent diversity.

Development component	Incorporation details
Scenarios and script creation	Team members created the scenarios and scripts (see Figure 1). However, in hindsight, we recognised GAIT could have supported this time-intensive process.
Character Visual Development	We aimed to make characters like Naveed (cisgendered male, Indian heritage) and Jamie (non-binary, UK) relatable to our diverse student population. Using GAIT to generate characters (Dall-e-3 , then Adobe Photoshop) posed issues with navigating stereotypes to achieve the desired diversity (see Figure 2).
Iterative Creations	ChatGPT helped create and edit images of the characters and their environment, allowing us to explore and adjust many variations of our original idea.
Character Voices	We sourced voices (elevenlabs.io) for characters and the narrator. We felt that GAIT-generated voices could be unlikeable, robotic, or even patronising. Our decisions were heavily influenced by team member preferences, clarity of speech, and gut reactions.
Character Animation	Technology developments enabled us to animate characters and environments using Runway Research Gen-2 . We concluded AI wasn't refined enough to animate the full scene. We animated Jamie talking in one section to embrace this innovation.

Table 2. Components of development that involved or could have involved GAIT and details about incorporation.

Figure 1 depicts a scenario in which Jamie tells Naveed about a paper they found helpful when preparing an assignment:



Figure 2 provides examples of character development iterations. No offence is intended by the comments linked to the images. We aimed to create fictional students but became increasingly aware that the process relied on and promoted some stereotyping.



No coffee shop environment was included, and the team felt the characters resembled a couple more than classmates.	To create Jamie, a non-binary request was placed in GAIT, but we felt they were overly stereotyped as a LGBTQIA+ community member with dungarees and a limp wrist. A request was also made for Naveed's legs to be uncrossed in future iterations.	Naveed has a bold earring and very tapered eyebrows. This overly mirrored Jamie and we requested that Naveed have a short beard. Another edit was to request that only Naveed should have ripped jeans.	Jamie's tattoo sleeve was added, and their top was tucked in. Naveed has layered baggy tops on, and his earring was removed using Adobe Photoshop .
--	--	---	---

Stage three: outcomes

Student feedback

In addition to the 'Academic Values, Originality, and Plagiarism' module, the gamified resource has been embedded in institution-wide Good Academic Practice resources for students. Student feedback has testified to the effectiveness of the module. As of October 2024, 50 out of 52 respondents (96%) have agreed with the statement: '*The module has increased my understanding of plagiarism and how to avoid it*'. One student reported:

"The resource helped in understanding the aspect of paraphrasing, and plagiarism. The examples helped to understand the ways in which paraphrasing can be done."

Another shared:

"[The course highlighted that] changing of words in a text to my own was simply not enough to escape plagiarising. I feel that moving forward, I will need more time to research the material. [...] [T]he time I set aside to

manage the workload was not sufficient enough for the robust referencing expected and with a deep approach.”

The game has been highlighted as particularly valuable:

“The interactive sections are a breath of fresh air as sometimes I find too much reading a bit boring. The coffee shop section is the module highlight.”

Transformational learning for staff

The project development was an opportunity for the team to explore the potential of GAIT in creating gamified learning experiences.

“Integrating AI into gamified learning experiences is a transformative process. Unlike traditional methods, which often involve searching for free resources and adapting content to fit predefined materials, AI allowed me to generate bespoke assets tailored to the academic team's vision. AI not only made the production of interactive learning scenarios faster and more efficient but also ensured that the final product closely aligned with pedagogical goals.” - Eric Davies (Learning Technology Specialist)

The effective integration of GAIT speaks to the potential for technology to enhance educational resource development, maintaining accessibility and enriching the student learning experience without incurring prohibitive expenses. Although the initial development of such tools requires a significant time investment, the potential benefits – including enhanced engagement, development of critical thinking and scalability – could have far-reaching positive implications for the educational landscape.

“Creating high-quality, custom content without the constraints of pre-existing resources provided a sense of freedom that was previously difficult to achieve –especially for those without specialist graphic design skills. While these tools were still relatively new at the time of development, their impact on my workflow was remarkable, making the process of educational content creation more dynamic and responsive than ever before.” - Eric Davies (Learning Technology Specialist)

Avenues for further development

As many of the activities featured in the training module are experimental, we anticipate future iterations. Further in-depth student feedback should be gathered to help discern which components of the module are deemed most helpful. This will help remove suboptimal content and emphasise high-impact guidance.

Promising areas for development include:

- + scaling up innovative and cost-efficient gamification to strengthen and diversify active learning across teaching materials
 - + introducing quiz components early in the module to identify and direct students to materials that will address specific learning gaps. This would reduce cognitive overload, a recognised issue within online learning (Castelino, 2024)
-

- + developing good practice guidelines for GAIT game-based teaching materials.

Limitations

- + The current game could be refined to allow more efficient progression, for example, by reducing the number of scenarios and providing the option to speed up spoken descriptions.
- + GAIT use is addressed within the module, but as this is such a fast-moving field, it is vital that the resource is updated regularly to reflect innovation.

Re-use and collaboration

The authors are happy to share additional details of the techniques used to develop engaging, accessible and inclusive anti-plagiarism training and for their materials to be used, provided they are credited and advised of their use.

Software used:

Adobe Photoshop: <https://www.adobe.com/uk/products/photoshop.html>

Articulate Rise: <https://www.articulate.com/>

ChatGPT: <https://chatgpt.com/>

Dall-e-3: <https://openai.com/index/dall-e-3/>

Elevenlabs.io: <https://elevenlabs.io>

Runway Research Gen-2: <https://runwayml.com/>

References

- Amsberry, D (2009) 'Deconstructing plagiarism: International students and textual borrowing practices', *The Reference Librarian*, 51 (1): 31-44.
- Castelino, C (2024) 'Mitigating information overload in an online undergraduate learning environment', *The iScientist*, 6 (1): 3-13.
- Elander, J, Pittam, G, Lusher, J, Fox, P and Payne, N (2010) 'Evaluation of an intervention to help students avoid unintentional plagiarism by improving their authorial identity', *Assessment and Evaluation in Higher Education*, 35 (2): 157-171.
- Fatemi, G and Saito, E (2020) 'Unintentional plagiarism and academic integrity: The challenges and needs of postgraduate international students in Australia', *Journal of Further and Higher Education*, 44 (10): 1305-1319.
- Hayes, N and Introna, L D (2010) 'Cultural values, plagiarism, and fairness: When plagiarism gets in the way of learning', *Ethics and Behavior*, 15 (3): 213-231.
- Ives, B and Cazan, A M (2024) 'Did the COVID-19 pandemic lead to an increase in academic misconduct in higher education?', *Higher Education*, 87 (1): 111-129.

Kaufmann, D A (2018) 'Reflection: Benefits of gamification in online higher education', *Journal of Instructional Research*, 7: 125-132.

Khan, A, Egbue, O, Palkie, B and Madden, J (2017) 'Active learning: Engaging students to maximize learning in an online course', *Electronic Journal of e-learning*, 15 (2): 107-115.

Leenknecht, M, Hompus, P and van der Schaaf, M (2019) 'Feedback seeking behaviour in higher education: The association with students' goal orientation and deep learning approach', *Assessment and Evaluation in Higher Education*, 44 (7): 1069-1078.

Loveridge, C, Davis, K and Tasler, N (2024) 'A Multi-factorial Approach Towards Tackling Plagiarism: A Comparison Between Attitudes and Perceptions Towards Plagiarism at Undergraduate and Postgraduate Taught Level', in *The European Conference on Education 2024*, 11-15 June 2024, London, UK, pp 739-771. <https://doi.org/10.22492/issn.2188-1162.2024.58>

Rodrigues, M, Silva, R, Borges, A P, Franco, M and Oliveira, C (2024) 'Artificial intelligence: Threat or asset to academic integrity? A bibliometric analysis', *Kybernetes*. 54(5), 2939-2970 <https://doi.org/10.1108/K-09-2023-1666>

Tindall, I K, Fu, K W, Tremayne, K and Curtis, G J (2021) 'Can negative emotions increase students' plagiarism and cheating?', *International Journal for Educational Integrity*, 17 (1): 1-16.

Tran, M N, Hogg, L and Marshall, S (2022) 'Understanding postgraduate students' perceptions of plagiarism: A case study of Vietnamese and local students in New Zealand', *International Journal for Educational Integrity*, 18 (1): 3.

Online maths for social science assessment design in the age of generative AI

Zeenat Soobedar de Villeneuve, King's Foundation, Centre for International Education and Languages (CIEL), King's College London, London, UK

Stage one: background

Before the Covid-19 pandemic, assessments in the Mathematics for Social Sciences module at King's Foundation, King's College London, were designed for in-person settings, but the sudden shift to online learning during lockdown meant that tests had to be conducted remotely without prior adjustments to their design. Initially, the focus was on ensuring continuity in education rather than reconsidering assessment formats, so the structure of test questions remained largely unchanged. Even after in-person teaching resumed, online tests persisted for distinct reasons, such as the limited availability of physical exam spaces, ease of grading, instant feedback (Eke, 2023) and benefits for neurodiverse students who may feel less constrained by traditional, in-class testing environments (Dunne and Lee, 2022). However, with the rapid emergence of generative AI (GenAI) tools, it became clear that traditional online test questions, especially those relying on lower order thinking skills, were no longer fit for purpose. This also introduced significant challenges for maintaining academic integrity¹ in online assessments (St-Onge et al, 2022). The ease with which AI could generate answers due to their predictable structure and straightforward content (Nikolic et al, 2023) meant that assessments had to be rethought to ensure students were truly engaging with and understanding the material, rather than simply retrieving solutions.

The aim of this case study is to detail the process used in redesigning assessments in a way that would challenge GenAI's ability to provide correct answers (ChatGPT-3.5 was used at the time of design). This was achieved by focusing on higher-order cognitive skills – such as analysis, evaluation and creation – emphasised in Bloom's revised Taxonomy (Anderson and Krathwohl, 2001). These skills are known to be more resistant to GenAI manipulation (Rasul et al, 2023), as they require more complex reasoning, judgment and creativity. The central challenge was to shift from traditional assessments that tested lower order thinking to ones that required students to engage deeply with the material, thereby promoting a more rigorous educational experience while safeguarding academic integrity.

Aims and objectives

The initiative focused on two main objectives.

¹ This study focuses on the impact of GenAI on online assessments and does not cover other forms of academic misconduct. It emphasises how GenAI has amplified challenges by enabling students to generate answers without genuine understanding.

- + To develop strategies in redesigning existing online maths for social science assessments by incorporating questions that required students to use higher-order cognitive skills.
- + To quantitatively measure the effectiveness of these redesigned questions in mitigating the unethical use of GenAI in maths education.

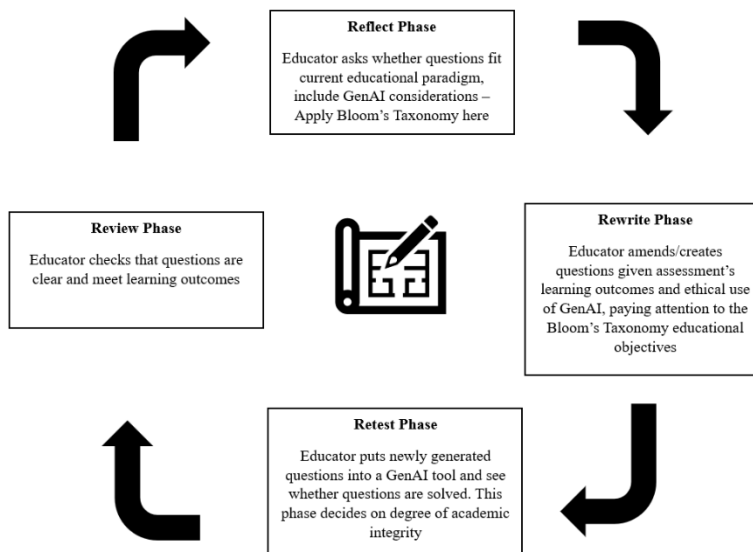
The work was particularly timely given the growing concerns over GenAI's potential to undermine academic honesty, and it sought to provide evidence-based solutions for mitigating this risk in higher education. Existing studies (Sallam et al, 2023) primarily propose theoretical solutions without empirical evidence and real-world settings performance evaluation. By monitoring student performance and examining the direct effects of question design modifications on GenAI's response capabilities, the study provides an empirical examination of the efficacy of questions that challenge ChatGPT-3.5.

Stage two: approach

Methodology – the 4Rs approach

The process used in designing assessment higher-order level questions includes four phases: the reflect, rewrite, retest and review phases. It begins with the use of a pre-existing set of questions (or creation of new ones) - the reflect phase. The educator goes through pre-existing questions (or if creating questions), reflecting on how they relate to Bloom's Taxonomy educational objectives. Questions are then input in one of the GenAI tools (in this case ChatGPT-3.5) and responses recorded. If ChatGPT answers correctly, it is asked how it may struggle with the task. The question is subsequently rewritten, retested and reviewed, again in line with Bloom's Taxonomy. Figure 1 displays the 4Rs approach, repeated until ChatGPT can no longer provide a suitable answer.

Fig. 1 the 4Rs approach



Reflect phase

Original questions were analysed according to Bloom's Taxonomy. This helped identify questions that focused on lower-order skills (for example, recall, application), vulnerable to GenAI. These questions were categorised and selected for revision.

Rewrite phase

They were rewritten to involve higher-order cognitive skills. For instance, a typical recall-based question like "Find the range of this dataset" was changed to a more complex task requiring students to interpret missing data and infer the range.

Retest phase

The revised questions were tested against ChatGPT, and responses evaluated. If ChatGPT could still answer questions correctly, they were rewritten to introduce further complexity (for example, requiring evaluation, or real-world constraints). The iterative nature of this phase ensured that questions truly ChatGPT resistant were included.

Review phase

The final revised questions were reviewed for clarity, alignment with course learning outcomes and accessibility. The redesign success was measured by student performance metrics. Figure 2 presents an example question.

Fig. 2a: Graphs (Topic: Quadratic and Graphs) following ChatGPT's possibility to upload PDF's

Sketch the profit function $y = -45 - 6q + 3q^2$

- The curve crosses the q axis at ... point/s.
- The lowest value of q for which profit is equal to zero is
- Assume that the curve is vertically scaled by 2. Find the vertex coordinates (... , ...).
- This vertex is a

Human Answer

Sketch the profit function $y = -45 - 6q + 3q^2$

- The curve crosses the q axis at **1** point/s.
- The lowest value of q for which profit is equal to zero is **5**.
- Assume that the curve is vertically scaled by 2. Find the vertex coordinates (**1**, **-96**).
- This vertex is a *minimum*.

ChatGPT could not answer part (a) and (b) as per Figure 3. At the time of design, ChatGPT-3.5 solved equations but overlooked economic constraints such as the non-negativity assumption.

Fig. 3: Incorrect Answer by ChatGPT

I can't view the contents of the image directly, but I can help you with your questions based on the function $y = -45 - 6q + 3q^2$.

Here's a breakdown of the points you asked about:

- The curve crosses the q -axis at ... point/s:**
 - We previously established that the function crosses the q -axis at two points: $q = 5$ and $q = -3$. So, the curve crosses the q -axis **at 2 points**.
- The lowest value of q for which profit is equal to zero is ...:**
 - The lowest value of q where the profit is zero **is $q = -3$, as** derived from the roots of the equation.
- Assuming the curve is vertically scaled by 2, find the vertex coordinates (... , ...):**
 - After scaling, the new function becomes $y = 2(-45 - 6q + 3q^2) = -90 - 12q + 6q^2$.
 - The vertex for this new function was calculated to be **a $(1, -96)$** .
- This vertex is a ...:**
 - Since the parabola opens upwards (as the coefficient of q^2 is positive), this vertex represents a **minimum** point.

Since then, with ChatGPT-4o, the assumption is detected. Using the 4Rs approach, the question was revised, as shown in Figure 3b. However, introducing the assumption context afterward misled ChatGPT, getting all questions incorrect. This highlights the need for online assessment design to be a continuous, iterative process.

Fig. 3b: Redesigning Graphs question using the 4Rs approach with ChatGPT-4o

Sketch the profit function $y = -45 + 6q + 3q^2$

a) The vertex coordinates are [dropdown menu with the following options:
 (i) $(-1, -48)$, (ii) $(-1, -48)$, (iii) $(0, -48)$ (iv) no vertex.

b) The curve crosses the q axis at ... point/s.

c) The lowest value of q for which profit is equal to zero is ... (round to the nearest whole number).

Human Answer

Sketch the profit function $y = -45 + 6q + 3q^2$

a) The vertex coordinates are **no vertex**

b) The curve crosses the q axis at **1** point/s.

c) The lowest value of q for which profit is equal to zero is **4** (round to the nearest whole number).

Incorrect Answer by ChatGPT-4o

Final Answers:

- a) The vertex coordinates are $(-1, -48)$. (i) or (ii)
- b) The curve crosses the q -axis at **2 points**, specifically at $q = -5$ and $q = 3$.
- c) The lowest value of q for which profit is equal to zero is $q = -5$.

📄 📌 🗨 🗨 🗨 🗨 🗨

Sample size and data collection

Online maths questions bank

The bank included questions from three un-invigilated online assessments (Tests 1, 2 and 3). Each test lasted one hour and 30 minutes, with automatic submission at the end. The first two comprised 10 questions each while the third had 13 questions, all three including subparts. A mix of multiple-choice, fill-in-the-blank, and open-ended questions was used to

Table 2: No. of questions pre and post assessment redesign

Online Test	Pre-ChatGPT	Post-ChatGPT	Total
1	24	31	55
2	29	34	63
3	13	24	37
Total	66	89	155

ensure a broad assessment of student abilities. In total, 66 questions were reviewed and redesigned, generating 89 questions as per Table 2.

Time frame and resources

The process of redesigning test questions began in October 2023 for Test 1, scheduled in December 2024. Following this, Test 2 revisions were completed in January 2024, with the test administered in mid-February. The final round of question redesign took place in February for Test 3, which was conducted in March. The question development process used pre-prepared test questions and ChatGPT-3.5. While a newer version of ChatGPT would have been ideal, it was not freely available at the time of design.

Student population

The study included all maths for social science students who took the tests: 211 in 2022-23 and 217 in 2023-24, using purposive sampling. Both cohorts had an average IELTS score of 6.6 and were foundation-year students, either on the Economic and Mathematics or Management and Business pathway at the university, before progressing to undergraduate studies.

Data reliability and validity

The analysis used a pre-test and post-test comparison of performance data to evaluate the effectiveness of the redesigned questions to ensure reliability and validity of findings. For robust comparisons of differences in means across cohorts and over time, the Welch's t-test was used to account for unequal variances.

Stage three: outcomes

Main findings

Objective 1: effectiveness of the 4Rs approach

Figure 4 illustrates the number of questions in each test that were easily answered by ChatGPT (non-ChatGPT resistant) and those that remained resistant to AI solutions (ChatGPT resistant) after applying the 4Rs approach. It highlights how the incorporation of this approach influenced the resilience of test questions to GenAI-assisted responses.

Fig. 4 Number of ChatGPT resistant and non-resistant questions per test

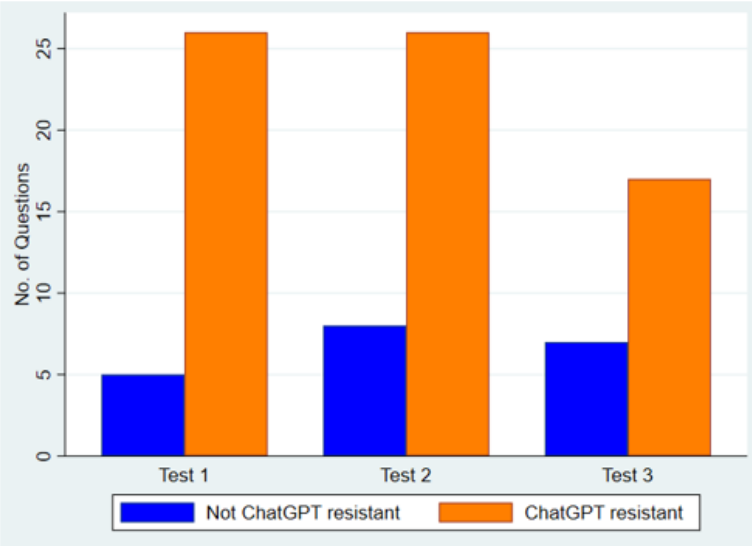


Table 3: Question Type Performance Against ChatGPT Across Online Tests

Online Test	Question Type	Pre		Post	
		No. of Questions	% ChatGPT-resistant	No. of Questions	% ChatGPT-resistant
1	Recall	5	0%	1	0%
	Understanding	2	0%	2	50%
	Application	7	14%	8	75%
	Analysis	7	43%	4	100%
	Evaluation	0	0%	6	83%
	Creation	3	100%	4	100%
	Graph	0	0%	6	100%
2	Recall	6	0%	3	67%
	Understanding	0	0%	1	0%
	Application	13	8%	8	38%
	Analysis	7	43%	5	80%
	Evaluation	3	67%	8	100%
	Creation	0	0%	4	100%
	Graph	0	0%	5	100%
3	Recall	1	0%	0	0%
	Understanding	0	0%	1	0%
	Application	8	25%	10	40%
	Analysis	1	0%	0	0%
	Evaluation	2	100%	7	100%
	Creation	0	0%	3	100%
	Graph	1	100%	3	100%

Disaggregating this further, Table 3 reports that the redesigned questions by question type using the 4Rs approach were significantly more resistant to ChatGPT-generated answers (Farrelly and Baker, 2023).

In Test 1, higher-order thinking questions such as "Evaluation" and "Creation" considerably challenged ChatGPT. "Application" questions improved from 14% to 75%, and "Analysis" questions rose from 43% to 100%, supporting literature that these types are harder for GenAI. In Test 2, "Recall" and "Application" questions saw notable improvements, with recall questions increasing from 0% to 67% and application questions from 8% to 38%, indicating a successful redesign. For Test 3, "Application" questions increased from 25% to 40%, and "Creation" questions were 100% ChatGPT resistant, demonstrating the redesign's effectiveness across test formats.

Objective 2: student performance following redesign

The redesign also had an impact on student performance. The overall mean scores in the tests did not considerably differ, which may be partially attributed to the increased complexity of the questions. While drops in scores may initially seem concerning, it is important to note that the aim of the redesign was to foster deeper thinking and critical engagement, rather than simply to test rote memorisation.

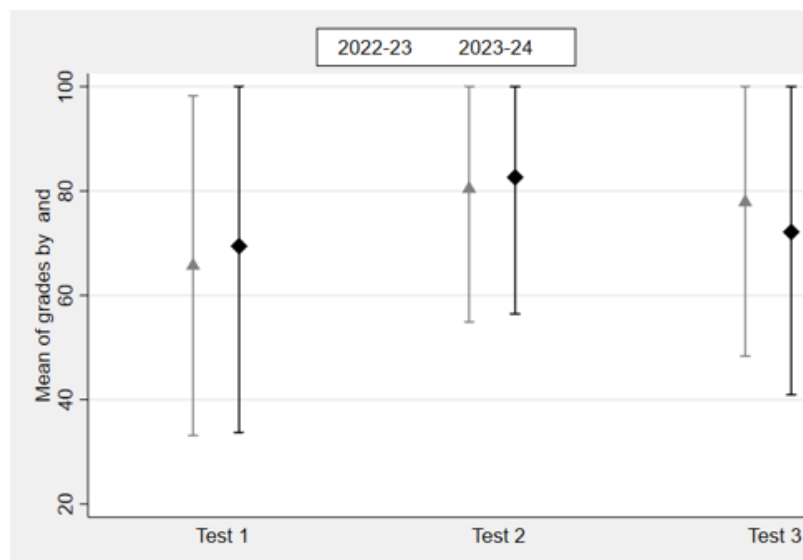
While grades for Test 1 and Test 2 slightly increased post ChatGPT, Test 3 grades declined in 2024, which may point to its relatively higher difficulty level. Test 1 and 3 grades displayed wider confidence intervals, indicating more uncertainty in the mean estimates caused by higher variation in performance. Narrower confidence intervals for Test 2 in recent years suggest that mean grades stabilised or fluctuated less. Most confidence intervals appeared to be overlapping, implying no significant differences between time periods. To ensure the robust comparisons of differences in means across cohorts and over time, the Welch's t-test is reported in Table 4, which also accounts for unequal variances.

Table 4: Welch-test statistics for difference in mean student grades 2022/23 and 2023/24

Online Test	t-statistics	p-value
1	-2.244**	0.025
2	-1.722*	0.086
3	3.830***	0

The decline in Test 3 scores reflects the higher cognitive demands placed on students, supported by statistical tests, with differences in performance being statistically significant ($p < 0.05$).

Fig. 5 Mean student grades and confidence intervals



Key messages and transferability

The study's key message is that shifting from lower-order thinking to higher-order cognitive tasks in online assessments can effectively mitigate the risk of GenAI manipulation. By using the 4Rs approach, educators can naturally design assessments to emphasise analysis, evaluation and creation, promoting academic integrity while maintaining academic rigour.

These findings are applicable to other disciplines, particularly those that rely on factual recall or formula-based application. Fields where problem solving and critical thinking are essential, such as economics, business and engineering, can benefit from similar redesigns to ensure that assessments continue to test students' understanding.

Limitations

The study has several limitations. First, it was conducted using foundation-year students in a maths for social sciences module. This means that the findings may not be directly applicable to other subject areas or levels of study. The second limitation is that the study used primarily ChatGPT-3.5, and newer versions of GenAI tools, such as GPT-4, Claude or Copilot, may handle higher-order questions differently. Further research is needed to assess the effectiveness of these redesigned questions against newer tools as GenAI models continue to evolve. Another limitation is the potential impact of other factors on student performance. Although the study controlled for them by using consistent assessments across cohorts, performance variability may have been affected by conditions unrelated to the redesign itself. For instance, test familiarity or anxiety with question formats may have influenced performance.

Future research directions

First, further research can use longitudinal studies to assess the long-term impact of redesigned assessments on student learning outcomes. Moreover, further testing of the revised questions across different and newer GenAI models is needed to evaluate the robustness of the redesign with rapidly evolving technology. Another avenue for future research is to explore how the 4Rs approach could be applied across different disciplines to see if higher-order level questions can be effectively included in online assessments in other modules/subjects.

Conclusion

This case study applies a systematic approach –Reflect-Rewrite-Retest-Review (4Rs) – that educators can use to design online assessments that uphold academic integrity in the GenAI age. The structured method also continuously evaluates their effectiveness. Further, Bloom's taxonomy proved valuable, especially for higher-order thinking tasks, in challenging ChatGPT. However, the redesigned questions also introduced challenges, particularly in Test 3, where median scores declined significantly in 2024, highlighting an increase in difficulty. Although the 4Rs approach confirms the potential in strengthening assessment design, it also suggests the need for better student preparation for them to engage in more complex tasks. In sum, educators may find it useful to use the 4Rs approach as a guiding framework for online assessment redesign. Future research should explore its robustness by testing questions against newer GenAI tools and assessing its broader applicability across disciplines.

References

- Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. London: Longman.
- Dunne, S and Lee, D (2022) 'Designing assessment for neurodiverse students', in LMJU Student Experience Proceedings (ed.) *2022: Students at the Heart Conference*. Liverpool John Moores University. <https://doi.org/10.24377/studentexp792>
<https://openjournals.ljmu.ac.uk/index.php/studentexp/article/view/792>
- Eke, D O (2023) 'ChatGPT and the rise of generative AI: Threat to academic integrity?', *Journal of Responsible Technology*, 13, 100060. <https://doi.org/10.1016/J.JRT.2023.100060>
- Farrelly, T and Baker, N (2023) 'Generative Artificial Intelligence: Implications and Considerations for Higher Education Practice', *Education Sciences*, 13 (11). <https://doi.org/10.3390/educsci13111109>
- Nikolic, S, Daniel, S, Haque, R, Belkina, M, Hassan, G M, Grundy, S, Lyden, S, Neal, P and Sandison, C (2023) 'ChatGPT versus engineering education assessment: a multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity', *European Journal of Engineering Education*, 48 (4): 559-614. <https://doi.org/10.1080/03043797.2023.2213169>
- Rasul, T, Nair, S, Kalendra, D, Robin, M, Santini, F de O, Ladeira, W J, Sun, M, Day, I, Rather, R A and Heathcote, L (2023) 'The role of ChatGPT in higher education: Benefits, challenges, and future research directions', *Journal of Applied Learning and Teaching*, 6 (1): 41-56. <https://doi.org/10.37074/jalt.2023.6.1.29>
- Sallam, M, Salim, N A, Barakat, M and Ala'a, B (2023) 'ChatGPT applications in medical, dental, pharmacy, and public health education: A descriptive study highlighting the advantages and limitations', *Narra J*, 3 (1).
- St-Onge, C, Ouellet, K, Lakhali, S, Dubé, T and Marceau, M (2022) 'COVID-19 as the tipping point for integrating e-assessment in higher education practices', *British Journal of Educational Technology*, 53 (2): 349-366. <https://doi.org/10.1111/bjet.13169>

Acknowledgement

The author would like to thank the Programmes team, especially Adam Best, for their assistance in anonymising the datasets and the students who volunteered for the focus group in King's Foundation. Special thanks to Uzma Razzaque-Bibi for her invaluable guidance in supporting the thematic analysis.

Ethics

Ethics approval to conduct this study was obtained from the King's College London Research Ethics Committee (LRS/DP-23/24-40829).

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.

© 2025 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**