

Reconsidering assessment for the ChatGPT era: QAA advice on developing sustainable assessment strategies

Scope and purpose of this advice

This paper sets out QAA's advice for providers on how to approach the assessment of students in a world where students have access to Generative Artificial Intelligence (AI) tools. The principles set out here are applicable to both higher and further education. This resource develops a theme first introduced in our earlier advice - [Maintaining quality and standards in the ChatGPT era: QAA advice on the opportunities and challenges posed by Generative Artificial Intelligence](#) - published in May 2023, around the (re)design of assessment strategies to mitigate the risks to academic integrity posed by the increased use of Generative Artificial Intelligence tools (such as ChatGPT) by students and learners.

The focus of this paper is on Generative Artificial Intelligence software that outputs text, but the principles may be applied to other Generative Artificial Intelligence tools - for instance, those that generate mathematical formulae, computer code images and other artifacts. For a brief introduction to this software, readers are directed to our earlier paper.

This paper includes:

- [Generative Artificial Intelligence as a catalyst for enhancing assessment strategy and practice](#)
- [selecting modes of assessment](#)
- [managing academic integrity](#)
- [creating a supportive environment](#)
- [a set of reflective questions](#)
- [table highlighting strengths and weaknesses of different assessment types](#)
- [additional resources](#)

Prompt for change

Awards are made, and classified, based on evidence that students have met the learning outcomes for their programme of study and achieved, or exceeded, certain threshold standards such as those articulated in [QAA's Subject Benchmark Statements](#). The rapid rise and ubiquity of Generative Artificial Intelligence software means that some or all the assessments that currently contribute to the evidence base may no longer be confidently ascribed to an individual student. This problem is compounded by the fact that the outputs of this software, despite its limitations, cannot reliably be detected, as explained in recent work by [Weber-Wulff and colleagues](#). This is something that will only become more challenging as the tools become more sophisticated and embedded in other software.



Generative Artificial Intelligence as a catalyst for enhancement

The rapid rise of Generative Artificial Intelligence software since OpenAI released ChatGPT in November 2022 has presented a range of challenges for multiple sectors. For the tertiary education sector, the focus has been on maintaining academic integrity, but it is also a powerful catalyst for change. Specifically, it offers a generational incentive for providers to require their programme and module teams to review and, where necessary, reimagine assessment strategies.

Reviewing assessment strategies

Three desirable outcomes of reviewing assessment strategies could be:

- Reducing the volume of assessment by removing items that are susceptible to misuse of Generative Artificial Intelligence tools to generate unauthorised outputs and repurposing the time available for other pedagogical activities.
- Promoting a shift towards greater use of synoptic assessments that test programme level outcomes by requiring students to synthesise knowledge from different parts of the programme. Some of these may permit or incorporate the use of Generative Artificial Intelligence tools.
- Developing a range of authentic assessments in which students are asked to use and apply their knowledge and competencies in real-life, often workplace related, settings. Ideally authentic assessments should have a synoptic element.

Creating space for Artificial Intelligence literacy

Reducing the volume of assessment creates space in the curriculum that can be used to develop a range of skills and competencies related to future employment, including how to use Generative Artificial Intelligence software in a discipline-appropriate fashion. Developing critical artificial intelligence literacy alongside other foundational academic skills, such as correct citation and referencing and critical thinking, is important for two reasons. First, because we anticipate that if it is not already a key graduate attribute, it very soon will be. Secondly, students will enter tertiary education with considerable prior experience of these tools without necessarily having the background to use them responsibly.

Aligning assessment and learning outcomes

In reviewing assessment strategies, programme teams need to ensure that assessment methods and criteria remain aligned to learning outcomes and teaching activities. This and other guiding principles for the update and (re)design of assessment strategies can be found in the Assessment section of the [UK Quality Code - Advice and Guidance](#). Mapping existing assessments to programme-level outcomes and removing any redundancy may be helpful in further reducing the assessment burden and identifying any gaps. Using an outcomes-based approach, formative and summative assessment enables student engagement and scaffolds learning to develop knowledge, competencies and skills. Further guidance on using an outcomes-based approach to teaching, learning and assessment can be found in the QAA [Membership Resources site](#) and in [Assessment in Digital and Blended Pedagogy](#) - the second of the QAA Hallmarks of Success Playbooks.

Choosing a range of assessments to support students and fit the programme

The precise selection of assessment types that are used by programme teams to allow students to evidence programme-level outcomes, while assuring academic integrity and meeting the requirements of professional, statutory and regulatory bodies will remain course and discipline specific. Some will choose to focus on forms of assessment that cannot currently be influenced by Generative Artificial Intelligence software, while others will transition to assessments that permit the use or actively incorporate generative tools in their completion. There is no single solution and the choices that programme teams make will be significantly influenced by institutional policies and the environment in which they are operating.

Selecting modes of assessment

The following range of assessment types is not intended to provide a comprehensive list of options but rather it can be used as a prompt for the types of assessment that could be deployed when developing programme-level assessment strategies.

Going backwards - unseen invigilated exams (handwritten)

One way in which the integrity of awards can be assured is to assess students exclusively or substantially using timebound unseen in-person invigilated examinations that require candidates to write their answers by hand. The strength of this traditional approach is that it allows the sampling of a range of learning outcomes, but at the same time it represents a regressive solution that would reverse much recent progress around accessibility. It is a form of assessment for which many cohorts of students are increasingly ill prepared in their prior education and one that is not authentic in that it tends to require a narrow range of competencies, such as the retention of facts as well as handwriting significant amounts of text under time limitations that are simply not relevant in many contemporary workplaces.

If you insist - unseen invigilated exams (digital)

There may be some, discipline-specific, scenarios in which time-bound examinations remain necessary and appropriate. In this circumstance it may be better to deliver examinations via a bespoke digitally secure platform. These allow unseen papers to be taken either in an invigilated examination hall on computers equipped with lockdown browsers or remotely, perhaps with the addition of digital proctoring software. It is a solution that, if required, allows online exams to remain as part of an assessment diet, while allowing word processed answers and preventing the use of Generative Artificial Intelligence software and other web-based resources. Further advice on online invigilation and other considerations around online assessment, including equity and student welfare, can be found in the QAA Membership publication [Digital Assessment Security](#).

The best of both worlds - observed examinations

An alternative approach is to observe a student complete one or more specific tasks related to their discipline or future employment and interview them about their understanding of the related principles, context and applications. Assessments can be designed to be both authentic and synoptic - for example, Observed Structured Clinical Examinations used in settings such as medical and nurse education in which candidates progress around a series of stations at each of which they are presented with a work-related task which they must complete and defend their approach. Candidates are marked on each task by multiple examiners according to a strict and objective rubric. This approach, which has been extended to the natural sciences using Observed Structured Practical Examinations, allows the efficient synoptic assessment of large numbers of candidates, although careful consideration needs to be given to scheduling to avoid unwanted sharing of questions between candidates.



Talking heads - oral examinations

With appropriate safeguards, oral (viva-voce) examinations also offer an opportunity to assess the breadth of a candidate's knowledge and understanding, including confirmation that the student was responsible for a written submission. Historically viva-voce examinations were used in some disciplines to assess candidates on classification borderlines of undergraduate degrees but have been discontinued at most providers on the grounds that they are unreliable and unnecessarily stressful to students. However, oral examinations in the form of structured interviews conducted by two or more examiners with clearly set out rubrics and appropriate safeguards for vulnerable students may be used formatively or summatively as a synoptic assessment. In addition, mini-vivas, in which small groups of students are interviewed together about their written submissions, can serve both to authenticate the work and contribute to its assessment. This approach, although resource intensive, is a powerful deterrent to students considering contract cheating. However, there are accessibility and inclusion factors that need to be considered with oral examinations - for example, recognising students with speech impairments and to ensure students with different accents are treated equitably.

Coursework that integrates Generative Artificial Intelligence by design

To help prepare students for their future workplace, it may be useful to design authentic coursework assessments which incorporate the use of Generative Artificial Intelligence software to complete routine and/or repetitive tasks and ask the students to reflect on its usefulness. Such an approach was the subject of a recent QAA Membership workshop - [Using Generative Artificial Intelligence for academic research](#) - which explored how novice undergraduate researchers can use a combination of ChatGPT and Google Scholar to complete a literature search quickly and reliably. This exercise demonstrates that the ability to use Generative Artificial Intelligence software ethically and effectively is an important competence for graduates to demonstrate and also provides an excellent example of learning through doing, which is a key tenant of competence-based education. Our recently published [Competence-based Education Primer](#) expands on this approach and will provide further inspiration for the design of authentic assessment.

Hybrid submissions and the future of academic writing

Developing discipline-relevant academic writing skills that allow learners to demonstrate their ability to acquire and synthesise knowledge remains an important element of most, if not all, degree programmes. Hybrid submissions of coursework that combine the output from Generative Artificial Intelligence tools with the learner's own work are already commonplace. In the short-term, allowing hybrid submissions in which the contribution of Generative Artificial Intelligence is fully acknowledged and is in keeping with institutional policies and guidelines, is a useful transitional arrangement as providers plan for the near future in which Generative Artificial Intelligence is embedded in the licensed software used by staff and students.

In planning for the future, it is vital to help learners progress from using Generative Artificial Intelligence software to complete simple short-form writing exercises to submitting a capstone dissertation, by providing ample formative opportunities to test their foundational academic skills, practise writing, and receive timely and constructive feedback on their outputs. Taking a long-term developmental approach avoids the immediate need to deploy unreliable detection software, retains the usefulness of written assessments and nurtures academic writing skills.

At a glance - strengths and weaknesses of different assessment types

Assessment type	Strengths	Weaknesses	Academic integrity	Sustainability
Invigilated unseen examinations (handwritten)	<p>Security - no (or only controlled) access to external physical or digital sources</p> <p>Synoptic - tests different learning outcomes via structure of paper and candidates' choice of questions</p> <p>Volume - can assess large numbers of students in parallel</p>	<p>Accessibility - challenging for students with certain characteristics to access plus rapidly decreasing number of students adequately prepared to handwrite large amounts of text</p> <p>Authentic - may only test a narrow range of knowledge / competencies</p> <p>Resources - places significant demands on a provider's estate</p>	High	Low
Invigilated unseen examinations (digital)	<p>Distribution - exams can be delivered offsite with appropriate digital security</p> <p>Security - with correct software, access to external physical or digital sources can be limited by the provider</p> <p>Synoptic - tests different learning outcomes via structure of paper and candidates' choice of questions</p> <p>Volume - can assess large numbers of students in parallel</p>	<p>Accessibility - there are challenges, but these may be easier to mitigate through technology</p> <p>Authentic - depending on assessment design, it may only test a narrow range of knowledge / competencies</p> <p>Security - even with digital proctoring, remote candidates can access digital assessment via other devices</p> <p>Resources - need for continual investment in digital security software</p>	Medium	Medium
Observed examinations	<p>Authentic - opportunity to apply competencies / knowledge to a range of realistic scenarios</p> <p>Synoptic - tests a wide range of competencies and understanding from different parts of the programme</p> <p>Volume - can assess significant numbers of students in parallel</p>	<p>Security - scheduling demands may mean that students assessed early in cycle can transmit information to those taking the assessment later</p>	High	High

<p>Oral examinations</p>	<p>Authentic - tests competencies that can be used in interviews, presentations and meetings</p> <p>Synoptic - tests a wide range of competencies and understanding from different parts of the programme</p>	<p>Accessibility - stressful for some students and challenging for those with certain characteristics (eg speech or hearing disabilities)</p> <p>Resources - consumes considerable amount of staff time</p> <p>Volume - suitable only for individuals or small groups at a time</p>	<p>High</p>	<p>Low</p>
<p>Coursework that integrates Generative Artificial Intelligence by design - use of these tools is part of the assessment brief and outputs are critiqued or reflected upon</p>	<p>Authentic - learning by doing, including using Generative Artificial Intelligence tools, can be built into assessment design</p> <p>Detection - detection is not necessary if using the AI tools is part of the assessment</p> <p>Resource - can repurpose existing assessments</p> <p>Synoptic - can be designed to test a range of knowledge / competencies and different elements of a module / programme</p>	<p>Accessibility - need to ensure fair access to Generative Artificial Intelligence tools for all students</p> <p>Resource - initial investment of time and ongoing review to take account of developments in AI tools.</p>	<p>High</p>	<p>High</p>
<p>Hybrid submissions - in which the use of Generative Artificial Intelligence tools are not part of the assessment brief (eg essays, dissertations)</p>	<p>Authentic - principally for those students who wish to continue to postgraduate education and beyond but also other careers</p> <p>Synoptic - tests the ability to synthesise knowledge and evidence from across a programme</p>	<p>Accessibility - need to ensure fair access to Generative Artificial Intelligence tools for all students</p> <p>Authentic - less obvious relevance for those exiting academia but still develops evaluation of evidence and synthesis</p> <p>Detection - heavily dependent on student declaration even at current state of evolution of tools and will be even more difficult when integrated in licensed software</p> <p>Resource - need to invest more resource in developing foundational academic skills on which to scaffold ethical use of Generative Artificial Intelligence</p>	<p>Low</p>	<p>Medium</p>

Managing academic integrity

In transitioning to new or updated patterns of assessment, it is important that providers create compassionate and supportive cultures in which students can serve as [champions of academic integrity](#) for their peers. Students and staff will inevitably make missteps on this journey as they adjust to a fast-changing environment and, in the first instance, academic misconduct policies should be applied sensitively and sparingly. Institutional policies increasingly recognise that submissions containing significant [unauthorised content generation](#) may not be regarded as [plagiarised using traditional definitions](#), but they do represent very poor academic practice as they do not represent the student's own work. In the first instance, suspected breaches of academic integrity guidelines may be best dealt with through the appropriate student support systems. However, in cases where an individual persistently exhibits poor academic practice through the inappropriate use of Generative Artificial Intelligence tools, they may be referred to their provider's academic misconduct procedures. Where a student is suspected of misconduct and has their case managed through academic misconduct processes, the level and extent of the perceived offence, along with any previous offences, should be considered. Some guiding principles can be found in the QAA Membership publication - [Academic Misconduct Penalties - Advice for providers](#).

Creating a supportive environment

The following conditions are likely to create a supportive environment for staff and students that can facilitate the changes to learning, teaching and assessment required by the advent of Generative Artificial Intelligence tools. These were shared in our [previous advice](#) and include having in place:

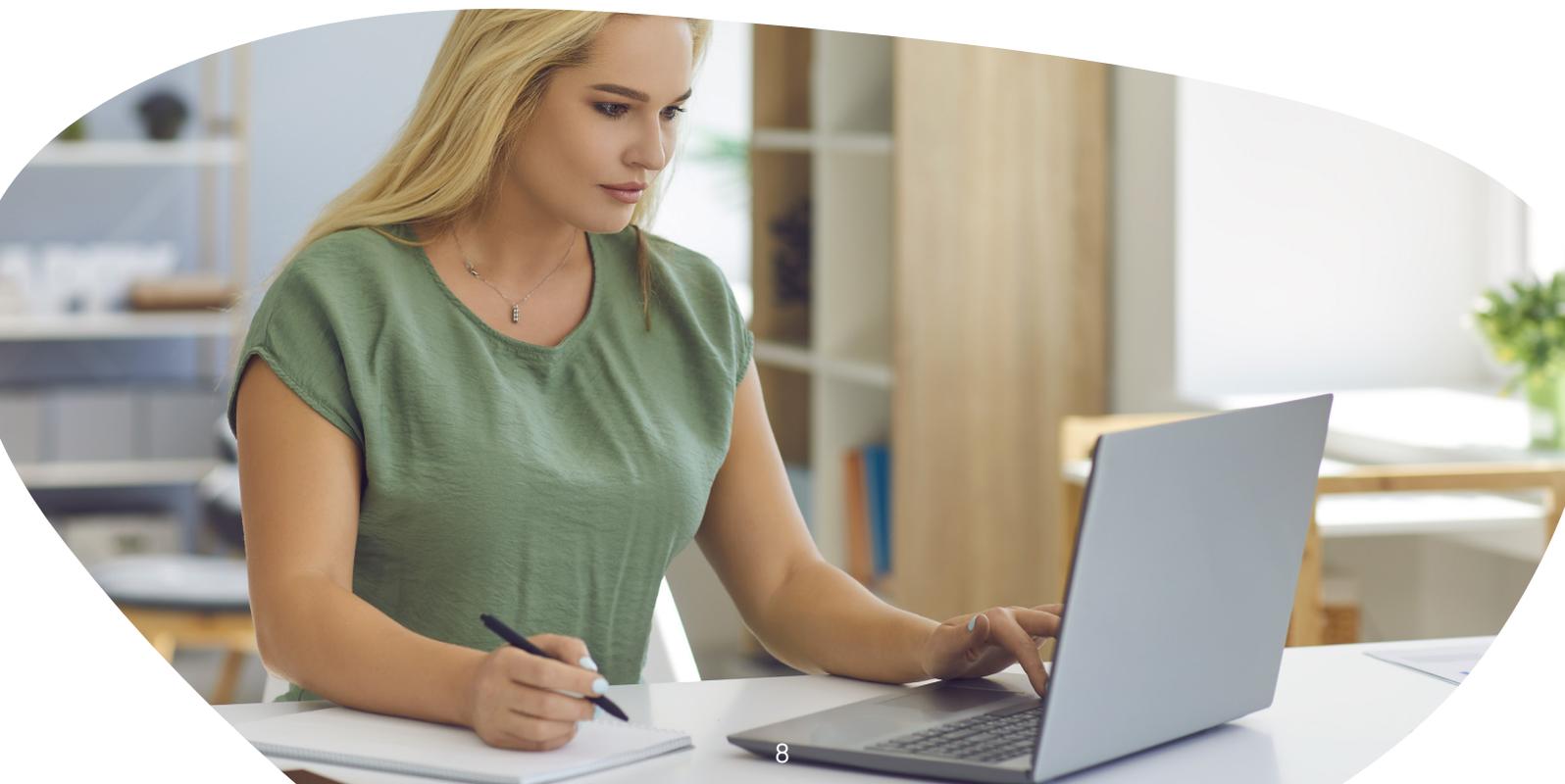
- Up-to-date policies and guidance around academic misconduct that explain how and in what circumstances Generative Artificial Intelligence tools may be used to prepare student submissions. In framing guidance for staff and students around the legitimate use of these tools, we encourage providers to reflect on the [Academic Integrity Charter for UK Higher Education](#) and the seven Principles for Academic Integrity it sets out. The guidance should be transparent, clearly communicated to all staff and students, and emphasise that academic misconduct is unacceptable and that responsibility for the integrity of the submission lies with the student.
- Opportunities for staff and students to learn and be updated regularly about other policies and frameworks that intersect with local academic misconduct guidelines; these might include policies around data protection and intellectual property.
- A digital literacy strategy - aims of which are clearly understood by staff and students - and which might be usefully informed by [Maha Bali's definition of critical artificial intelligence literacy](#), an ability to know how to use artificial intelligence tools, when it is appropriate to do so and how to critique the credibility and accuracy of their output. Implementation of such a strategy requires academic and professional staff to be familiar with institutional policies and have opportunities to learn how artificial intelligence tools function and the ways in which assessments can benefit from redesign to ensure the academic integrity of awards.
- Simple and fair arrangements for students to access the Generative Artificial Intelligence tools they need to complete their submissions. These policies need to ensure, as far as possible, that no advantage can be accrued by students who can access more advanced Generative Artificial Intelligence tools from behind a paywall. The policies are likely to need regular updating to take account of the future integration of tools into software packages that are licensed by the provider for student use.

The importance of remaining vigilant and agile

The advice in this paper is current at the time of writing (July 2023) and can be adapted by providers to suit their mission and student cohorts to support the review of existing assessment practices. However, the technology is evolving rapidly and providers will need to review and update their policies more frequently in response to the rate at which artificial intelligence software is developing and launching new products. It is essential to tell students where they can find the most up-to-date guidance. This can be supported by adding references to the location of current institutional policies on Generative Artificial Intelligence tools and academic integrity to course materials - such as module handbooks and assessment briefs.

Reflective questions for providers and programme teams

1. Have your educational strategies, student support mechanisms and academic misconduct processes been reviewed, updated and appropriately resourced to help students navigate this significant change in the way they learn, are taught and assessed?
2. How do the programme-level assessment approaches link to provider-level strategies? Is this clear for students, especially those studying programmes across more than one department (for example, joint degrees)?
3. Do your provider-level academic misconduct regulations include unauthorised, undeclared or otherwise inappropriate use of Generative Artificial Intelligence tools and the consequent outcomes? How do the provider-level regulations relate to programme-level guidance to students? Is the language clear and accessible for students and all staff?
4. Has your review of the existing assessment pattern for the programme identified any duplication, gaps or items that might be vulnerable to the use of Generative Artificial Intelligence tools and, after revision, do the proposed assessment methods and criteria remain aligned to learning outcomes and teaching activities?
5. Does your approach to assessment provide the best and most appropriate way for students to demonstrate their learning and key competencies, including the use of Generative Artificial Intelligence software, in a way that is sustainable, inclusive and assures academic integrity?
6. Do your arrangements for students submitting assessed work include a declaration of authenticity that considers the responsible use of artificial intelligence tools?



Additional resources

From QAA:

- Our [Chat GPT and Artificial Intelligence](#) webpage hosts links to a range of QAA and external resources relating to Generative Artificial Intelligence software.
- QAA Membership has supported two Collaborative Enhancement Projects relating to academic integrity:
 - [Accessibility and Equity in Proofreading](#) led by Loughborough University
 - [The improvement of student learning by linking inclusion / accessibility and academic integrity](#) led by Oxford Brookes University.
- Curated outputs from the QAA funded project [Learning from the Online Pivot: Approaches to aligning assessment and learning outcomes for student success](#) which QAA commissioned from Professor Elizabeth Cleaver and Professor Mike McLinden, independent higher education consultants.

External resources:

- A contribution by Professor David Boud (Deakin University) [Positioning assessment differently in a world of gen AI](#) to the webinar *Artificial Intelligence and Pedagogy: Advancing Personalized Learning, Adaptive Teaching, and Values-Based Assessment*, recently hosted by the University of Kent.
- The Russell Group has recently published its [New principles on the use of AI in education](#).
- The most recent European Network for Academic Integrity (ENAI) [Recommendations on the ethical use of Artificial Intelligence in Education](#) focus on the importance of equipping stakeholders with the skills and knowledge to use Generative Artificial Intelligence tools ethically and the need to develop and implement relevant educational policies addressing the opportunities and challenges posed by artificial intelligence in education.
- The [Programme Assessment Strategies \(PASS\)](#) project hosted by the University of Bradford offers further information about the design of an effective, efficient, inclusive and sustainable assessment strategy which delivers the key programme outcomes.
- [Programme design and delivery through the lens of academic integrity](#) - a short webinar hosted by Quality and Qualifications Ireland and presented by Kane Murdoch, Manager, Conduct and Integrity Unit, University of New South Wales, which considers some issues around programme design, delivery and assessment, and academic integrity.

Opportunities to discuss the use of Generative Artificial Intelligence tools in tertiary education

Keep in touch with our events by checking our [Generative Artificial Intelligence webpage](#) and sign up for weekly news with our [QAA Member update](#).

QAA advice and guidance is funded through membership fees, and is an expression of the sector's shared commitment to protecting academic standards and enhancing quality.

Published - July 2023

© The Quality Assurance Agency for Higher Education 2023
Registered charity numbers 1062746 and SC037786
www.qaa.ac.uk