

**Designing for the AI Era: A Framework for
Integrating Generative AI into Assessment in
Higher Education**

Practitioner Research
In Higher Education
Copyright © 2026
University of Cumbria
Online First Pages 51-70

Emma Ransome
Birmingham City University

Abstract

The rapid emergence of generative artificial intelligence (GenAI) presents both significant challenges and opportunities for assessment in higher education. Many existing assessment designs, developed prior to widespread AI use, are increasingly misaligned with the skills, practices, and forms of judgement required in an AI-enabled world. At the same time, sector debates around academic integrity, equity, and responsible use highlight the need for clear pedagogic guidance for staff and students.

This paper introduces the GenAI-Integrated Assessment Design Framework, a conceptual model to support the design of robust, authentic, inclusive, and future-facing assessment. Drawing on assessment theory, emerging GenAI scholarship, and principles of equity and transparency, the framework articulates five interdependent dimensions: alignment of assessment purpose and learning outcomes; authenticity and professional relevance; transparent expectations for AI use; diversification of assessment modes; and attention to integrity, equity, and accessibility.

The paper outlines how the framework can be operationalised at module, programme, and institutional levels, supporting curriculum design, quality assurance, and policy development beyond risk-driven responses.

Introduction

The rapid emergence of generative artificial intelligence (GenAI) has disrupted long-established assumptions about learning, teaching and assessment in higher education. Large language models and related tools are now embedded in students' everyday study practices, supporting activities such as explaining concepts, summarising readings, generating text and offering feedback on drafts. Recent survey data from the UK Higher Education Policy Institute show that 92 per cent of undergraduates' report using AI tools in some form, and 88 per cent have used GenAI specifically for assessments, a dramatic increase from 2024 figures (Freeman, 2025). This scale and speed of adoption has intensified long-standing debates about academic integrity, the credibility of awards and the kinds of skills graduates require in increasingly AI-mediated workplaces.

Policy and regulatory bodies have begun to respond. UNESCO's first global guidance on GenAI in education calls for "human-centred" approaches that protect human agency, equity, cultural and linguistic diversity, and academic freedom, while recognising that GenAI will become a pervasive feature of educational and professional life (UNESCO, 2023). In the UK, the Office for Students has framed AI as a significant source of both opportunity and risk, highlighting its potential impact on academic integrity, the meaning of assessment outcomes and patterns of access and participation (Office for Students, 2025). Jisc's guidance for college and university governors similarly emphasises the dual nature of GenAI: it can improve productivity and enrich learning, but also raises concerns around privacy, data security and fairness, and therefore demands institution-wide approaches to responsible adoption (Attewell, 2024).

Citation

Ransome, E. (2026) 'Designing for the AI Era: A Framework for Integrating Generative AI into Assessment in Higher Education', *PRHE Journal Online First*, pp. 51-70.

At institutional level, emerging research suggests that universities are moving from prohibition towards more nuanced positions that seek to balance innovation with academic standards. In a global analysis of adoption policies across 40 universities in six regions, found that institutional guidelines commonly stress academic integrity, enhancement of teaching and learning, and equity of access, but remain uneven and fragmented in their practical implications for assessment design (Jin et al., 2025). Within the UK, sector bodies such as the Quality Assurance Agency (QAA) have issued advice and resources on GenAI, including webinars and collaborative enhancement projects that explicitly ask what assessment should look like in the era of ChatGPT and how to reconceptualise “authentic assessment” under these conditions (QAA, 2023). These developments reflect a growing recognition that assessment cannot simply be insulated from GenAI; instead, assessment strategies need to be “stress-tested” against the capabilities of AI and redesigned where necessary to remain valid and meaningful (Freeman, 2025).

Parallel to policy debates, a rapidly expanding research literature has begun to interrogate how GenAI affects academic integrity, student learning behaviours and staff practice. Bittle and El-Gayar’s (2025) systematic review of GenAI and academic integrity in higher education concludes that GenAI both complicates traditional notions of plagiarism and contract cheating and opens up possibilities for formative feedback, scaffolding and personalised support when used transparently and ethically. Kofinas, Tsay and Pike (2025) focus specifically on authentic assessment, demonstrating that although such assessments have often been promoted as less vulnerable to academic misconduct, GenAI tools can still undermine their integrity if tasks are not carefully designed and if expectations about AI use are unclear. Together, these studies suggest that naive reliance on either traditional exams or unmodified ‘authentic’ tasks is insufficient; a more fundamental rethinking of assessment design is required.

At the same time, equity and inclusion concerns are increasingly visible in the literature and in policy. UNESCO’s guidance highlights the risk that GenAI could exacerbate existing inequalities, particularly where access to high-quality tools or AI literacy is uneven (UNESCO, 2023). The HEPI–Kortext survey shows ‘persistent digital divides’ in AI use and confidence, with students from more advantaged backgrounds and male students more likely to be frequent users (Freeman, 2025). QAA’s transnational project on ethical uses of GenAI in assessment has similarly reported uncertainty among students about what constitutes responsible use and how to balance convenience with moral and academic obligations (QAA, 2025). These findings raise critical questions about how assessment design can both uphold academic standards and ensure that expectations around GenAI are transparent, teachable and equitable.

Despite this emerging body of work, there remains a notable gap in the literature: institutions and programme teams lack clear, theoretically informed and practically usable frameworks for integrating GenAI into assessment design. Existing publications often focus either on high-level policy recommendations or on specific case studies of AI use in particular disciplines (Halliday, 2025; Francis, 2025), rather than providing a holistic model that links assessment purposes, task design, AI-use expectations, and equity and integrity considerations. Moreover, many institutional policies emphasise what students must *not* do with AI, without offering parallel guidance on how assessments can harness AI productively to develop critical judgement, ethical decision-making and AI literacy as graduate attributes.

This conceptual paper responds to these gaps by proposing the GenAI-Integrated Assessment Design Framework, a model intended to support educators, programme leaders and quality assurance teams in re-designing assessments for an AI-enabled higher education landscape. The framework is informed by three strands of literature:

- Scholarship on assessment for learning, authentic assessment and programme-level assessment strategies
- Emerging research on GenAI, academic integrity and AI literacy in higher education and
- Policy and guidance documents from international organisations, national regulators and sector bodies addressing responsible AI use

It conceptualises assessment design across five interdependent dimensions:

- Aligning assessment purpose with learning outcomes
- Designing for authenticity and professional relevance
- Establishing transparent expectations for AI use
- Reconsider assessment modes and evidence types and
- Ensuring integrity, equity, and accessibility.

The aim of the paper is twofold, firstly, it offers a theoretical contribution by articulating a coherent framework that brings together pedagogic, ethical and policy perspectives on GenAI and assessment. Secondly, it provides a practical contribution by illustrating how the framework can inform module and programme assessment strategies, staff development and institutional policy. By reframing GenAI as a catalyst for thoughtful assessment innovation rather than solely a threat to be contained, the paper seeks to support higher education providers in developing assessment practices that are robust, inclusive and aligned with the realities of AI-rich academic and professional environments.

Literature Review

Student engagement with generative AI in UK higher education

Since late 2022, student use of generative AI tools has rapidly normalised across UK higher education. Drawing on a 2024 poll of 1,250 undergraduates via UCAS, Freeman (2024) reports that over half of UK students (53%) had already used GenAI to help with assessments, most commonly as an “AI private tutor” to explain concepts, while only a small minority indicated deliberate use for cheating. A follow-up HEPI survey (Freeman, 2025) found that 92% of undergraduates were using AI tools and 88% had used GenAI specifically in relation to assessment, with time-saving and perceived improvements in assignment quality cited as key drivers.

These findings are broadly consistent with Jisc’s *Student Perceptions of AI 2025* report, which synthesises conversations with students across UK institutions. Jisc notes that students increasingly regard AI as part of their everyday study toolkit, but also express concerns about unclear institutional rules, uneven staff guidance and a desire for explicit teaching about how to use AI ethically and effectively. Importantly, both HEPI and Jisc identify disparities in usage and confidence: students from more advantaged backgrounds and male students are more likely to be frequent and confident users of AI tools, raising questions about digital inequality and differential access to the benefits of GenAI.

Taken together, these UK data suggest that generative AI is now deeply embedded in student assessment practices, often in ways that blur the distinction between learning support and potential misconduct. The literature consistently highlights a mismatch between students’ routine use of AI and the limited, inconsistent guidance many receive from their institutions, creating a pressing need for clearer expectations and assessment designs that “assume AI” rather than ignore it.

Sector-level policy and guidance on GenAI and assessment

Sector bodies and regulators in the UK and internationally have begun to respond to these developments. UNESCO’s global guidance on generative AI in education frames GenAI as a technology that demands a “human-centred” response, emphasising the protection of human agency, equity, diversity and academic freedom while encouraging thoughtful integration into education systems

(UNESCO, 2023). In the UK, Jisc's (2023) *Artificial Intelligence in Tertiary Education* report maps emerging use cases and highlights ethical considerations, including transparency, data privacy and the need for institutional strategies that go beyond ad-hoc experimentation.

The Quality Assurance Agency (QAA) has produced some of the most explicit UK guidance on assessment in the GenAI era. Its paper *Reconsidering Assessment for the ChatGPT Era* urges providers to review the validity and robustness of existing assessment portfolios in light of GenAI capabilities, to reconsider assessment types, and to focus on designing tasks that emphasise application, judgement and process rather than solely text production (QAA, 2023). The document stresses that “business-as-usual” assessment strategies are unlikely to remain sustainable and advocates for proactive redesign rather than reliance on detection tools.

From a regulatory perspective, the Office for Students (OfS) has signalled that AI is both an opportunity and a risk within its quality and standards remit. In a 2025 blog, the OfS outlines a principle-based approach to AI, encouraging providers to harness AI where it enhances student outcomes and labour market readiness, while safeguarding academic standards and protecting students from harm. Although not prescriptive about specific assessment forms, the OfS highlights the importance of institutions understanding AI's impact on assessment reliability, the credibility of qualifications and equality of opportunity.

Collectively, these sources indicate a sector-level recognition that assessment design must adapt to AI. However, they tend to provide high-level principles rather than granular tools to support module and programme teams in redesigning specific assessment tasks.

Generative AI, assessment and academic integrity

The relationship between GenAI and academic integrity has been a central concern in both research and media coverage. Adams (2024) reports on studies such as the University of Reading experiment—in which AI-generated exam scripts were submitted undetected and often outperformed genuine student work—have heightened fears that traditional, text-based assessments are increasingly vulnerable to AI “spoofing” and that detection is unreliable. Analyses of the HEPI data have prompted commentators and sector bodies to call for universities to “stress-test” existing assessments against AI capabilities rather than rely primarily on punitive policies or detection technologies (Foster and Borrett, 2025).

At a scholarly level, Luo's (2024) *A Critical Review of GenAI Policies in Higher Education Assessment* synthesises institutional responses to GenAI across a range of universities and argues that many policies remain narrowly focused on preserving a traditional notion of “originality” as the absence of external machine assistance. Luo contends that this conception is increasingly untenable given the ubiquity of AI and risks marginalising students who use GenAI transparently for legitimate learning purposes. Instead, the article calls for a reframing of originality towards “intellectual contribution” and critical engagement, including where AI tools are used as part of the workflow.

Sallai et al. (2024), drawing on research at the London School of Economics, likewise caution that “mindless” use of GenAI in assignments can bypass learning, but argue that prohibition is neither realistic nor educationally desirable. They propose that universities should approach GenAI “proactively”, integrating it into curricula and assessment in ways that foreground metacognition, critical reflection and the development of AI literacy, rather than treating it solely as a threat to be managed.

These studies indicate that the integrity challenge in the UK context is not simply one of detecting misconduct, but of clarifying how GenAI can legitimately contribute to learning and assessment while maintaining meaningful standards.

Emerging frameworks for AI-integrated assessment

A growing but still limited body of work provides more concrete frameworks for integrating GenAI into assessment design. Perkins et al.'s (2024) *Artificial Intelligence Assessment Scale (AIAS)* has been particularly influential. Developed initially through collaboration between institutions in Vietnam, Australia and the UK, the AIAS proposes a five-level scale, from “No AI” to “Full AI”, which enables educators to specify the permissible extent of AI involvement in each assessment task and to align this with intended learning outcomes. A subsequent pilot study at British University Vietnam, reported by Furze et al. (2024), found that using the AIAS was associated with reduced academic misconduct cases and modest gains in attainment and pass rates, suggesting that structured, transparent integration of AI may support both integrity and performance.

Although the AIAS emerged outside the UK, it has begun to be adopted by UK universities as part of their assessment guidance. Canterbury Christ Church University (2025), for example, has embedded the AIAS into its module handbook templates to help staff decide and communicate appropriate levels of GenAI engagement in assessment, positioning it as a tool for alignment between learning outcomes and AI use expectations. Projects at the University of Warwick similarly draw on AIAS principles in exploring GenAI-supported assessment, reinforcing the framework's relevance to UK HE practice.

More explicitly UK-based conceptual work includes Oc et al.'s (2025) *Contextual Assessment Design in the Age of Generative AI*, which argues that assessment decisions must be situated within disciplinary, institutional and regulatory contexts rather than framed as universal technical fixes. Oc et al. position GenAI as both a disruptor of established assessment forms and a catalyst for renewing attention to authentic, situated tasks that reflect real-world professional practice. The paper emphasises the need for “contextual judgement” by educators when determining whether and how GenAI should be incorporated into particular assessments, highlighting tensions between innovation, risk management and resource constraints.

Complementing these conceptual contributions, Oc et al.'s (2025) study of *Generative AI in Higher Education Assessments* examines how perceived risk and students' self-assessed “tech-savviness” influence their intention to use GenAI in assessment contexts. Their findings suggest that students who feel more confident with technology and perceive lower risks are more likely to adopt GenAI tools, underscoring the importance of institutional efforts to develop AI literacy and clarify acceptable use, especially for students who may be more cautious or marginalised.

Internationally, Xia et al.'s (2024) scoping review of GenAI in higher education, highlights a proliferation of small-scale implementations, position papers and ethical debates, but notes a relative scarcity of comprehensive, theoretically grounded assessment frameworks that can be readily applied across programmes. Together, these studies signal a shift from initial alarm towards more constructive attempts to articulate how AI can be integrated into assessment design; however, they remain fragmented and often focus on either policy tools (such as AIAS) or high-level conceptual reflections, rather than providing a holistic model that combines assessment purpose, task design, AI-use expectations and equity considerations.

Equity, inclusion and AI literacy in assessment

Equity and inclusion concerns run through both policy documents and empirical studies. UNESCO's (2023) guidance explicitly warns that GenAI may exacerbate existing inequalities, particularly where access to high-quality tools, digital infrastructure and AI literacy is uneven, and calls for proactive strategies to avoid deepening digital divides. In the UK, HEPI's 2025 survey reports significant demographic differences in AI use, with male students and those from higher socio-economic

backgrounds more likely to be frequent and confident users, echoing broader patterns of digital capital.

Jisc (2025) likewise notes that students with limited prior experience of AI often feel anxious and unsure about expectations, especially when institutional messaging focuses on misconduct rather than guidance. Students interviewed for the report express a desire for structured teaching on AI, including worked examples of disciplinary-relevant uses and explicit discussion of where AI use is appropriate within assessments. These findings are echoed in UK case studies and commentaries, which suggest that “hidden curricula” around AI—where more digitally confident students develop sophisticated AI workflows while others rely on ad-hoc experimentation or avoid AI altogether—risk creating new forms of attainment gap if left unaddressed.

From an assessment perspective, this raises two critical issues. First, without clear, scaffolded expectations, students may inadvertently breach academic integrity policies or under-utilise tools that could support their learning. Second, if assessments tacitly assume access to and fluency with AI tools without providing explicit development opportunities, they may privilege students with existing digital capital. The literature therefore points to the need for assessment frameworks that foreground transparency of AI expectations and embed AI literacy as an explicit learning outcome, rather than treating AI use as an optional or peripheral skill.

Assessment purpose, subjectification and student agency

Alongside the policy and integrity debates outlined above, a broader philosophical question concerns the purpose of assessment itself. Biesta (2009) distinguishes three interrelated functions, or domains of educational purpose, namely qualification, socialisation and subjectification. Qualification refers to the development of knowledge, skills and dispositions that enable individuals to participate in professional, social and civic life. Socialisation concerns induction into existing cultural, disciplinary and societal traditions, norms and practices. Subjectification refers to the process through which learners come to exist as autonomous subjects, developing agency, independence and the capacity for critical judgement and action. Biesta argues that contemporary education systems have increasingly prioritised measurable outcomes and standardised performance, often leaving limited space for students to develop their own voice, perspective and agency.

This tension is particularly relevant within AI mediated assessment environments. Early institutional responses to generative AI have frequently prioritised academic integrity concerns, often through increasingly prescriptive guidance, restrictive policies and attempts to regulate student use (Luo, 2024; QAA, 2023). While such approaches may reduce ambiguity and support consistency, they also risk narrowing assessment practices towards tightly specified outcomes and procedural compliance. Scholars have cautioned that responses centred primarily on control and risk management may inadvertently undermine opportunities for critical engagement and meaningful learning (Sallai et al., 2024; Jisc, 2025). Authentic assessment may provide one means of preserving space for subjectification by creating opportunities for contextual judgement, interpretation and engagement with meaningful problems. The challenge posed by generative AI is therefore not solely technical or integrity related. It also invites a broader reconsideration of whether assessment enables students to participate as active meaning makers rather than producers of compliant outputs.

Developing this capacity for judgement also raises questions regarding how students come to understand quality and standards in assessment. Research suggests that transparency through written criteria or rubrics alone is often insufficient. Rather than acquiring standards through documentation, students and tutors develop shared understandings through dialogue and critical engagement with exemplars (Rust, Price and O'Donovan, 2003; O'Connell et al., 2016). Bearman and Ajjawi (2021) similarly argue that calibration activities support the development of evaluative judgement by enabling

participants to discuss, compare and interpret examples of work. In AI mediated contexts, where students increasingly engage with AI generated outputs, opportunities to evaluate and critique exemplars may become especially important. Such activities may strengthen both assessment literacy and the evaluative agency associated with subjectification.

Synthesis and identified gap

Across UK policy reports, institutional guidance and emerging empirical and conceptual research, several themes converge. There is broad consensus that:

- Generative AI has become a routine feature of students' assessment practices in UK higher education (Freeman, 2025).
- Existing assessment portfolios, particularly text-heavy, decontextualised tasks, are increasingly vulnerable to undetected AI assistance and may no longer provide valid evidence of individual learning (QAA, 2023).
- Sector bodies advocate for more authentic, diversified assessment and for clearer institutional expectations regarding AI use but offer limited operational tools at programme or module level (OfS, 2025; QAA, 2025).
- Frameworks such as the AI Assessment Scale (AIAS) and contextual design approaches provide valuable starting points for thinking about AI-integrated assessment yet tend to focus either on categorising AI use or on high-level principles rather than offering a comprehensive design model (Perkins et al, 2024).
- Equity and AI literacy are central concerns, with evidence of differential access to and confidence with AI tools across student groups (Freeman, 2025).

However, the literature also reveals a significant gap; programme leaders and assessment designers lack an integrated, theoretically informed framework that:

- Begins with assessment purpose and learning outcomes in an AI-mediated environment
- Connects these to concrete decisions about task authenticity, modes of evidence and the permitted, prohibited or required use of GenAI
- Embeds transparency, equity and integrity considerations throughout the design process.

The remainder of this paper addresses this gap by proposing the GenAI-Integrated Assessment Design Framework, which synthesises insights from UK and international scholarship, policy and practice into a structured model intended to support coherent, context-sensitive assessment design in an AI-enabled higher education landscape.

Ethical Use of AI in Developing This Conceptual Framework

The development of this conceptual framework involved the transparent and ethical use of generative AI as a scholarly support tool, aligned with sector guidance on responsible academic practice. AI was used to assist with literature discovery, summarisation and organisation, enabling a broad scan of recent UK-based policy reports, sector guidance and empirical studies published between 2020 and 2025. In line with QAA (2023) and UNESCO (2023) principles, AI was not used to generate arguments, interpretations or conceptual claims autonomously; rather, it supported the author's established academic expertise in assessment design, artificial intelligence in education, and quality enhancement. All analytical decisions, theoretical development and final wording of the framework were undertaken by the author, with AI providing a means to improve efficiency, extend the scope of the review, and reduce mechanical workload.

This approach reflects the broader pedagogic stance underpinning the paper: that generative AI can enhance human scholarly work when used critically, transparently and accountably, but must not replace intellectual judgement or authorship. Consistent with emerging norms of responsible AI

scholarship, the author-maintained oversight of all sources, verified the accuracy and credibility of policy documents and research evidence, and ensured that the framework reflects human-led evaluation rather than machine-generated synthesis. By adopting this reflective and principled use of AI, the paper models the kind of ethical academic practice that higher education institutions may wish to encourage in staff and students as they navigate AI-supported learning and assessment.

The GenAI-Integrated Assessment Design Framework

The GenAI-Integrated Assessment Design Framework is proposed as a structured, theoretically informed model to support educators, programme leaders and quality teams in redesigning assessment for an AI-enabled higher education environment. Developed through the synthesis of recent UK policy guidance, empirical studies, and contemporary assessment theory, the framework is intended to provide a practical yet conceptually robust tool for navigating the complex interplay between generative AI, academic standards, equity, and authenticity. The model as illustrated in Figure 1. comprises five interdependent dimensions, each addressing a critical aspect of assessment design in the context of GenAI.

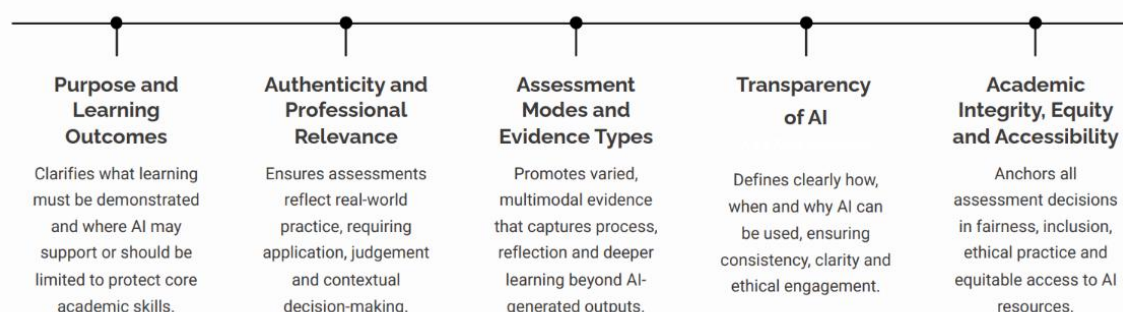


Figure 1. The GenAI-Integrated Assessment Design Framework.

Together, these dimensions guide educators through the full lifecycle of assessment design, ensuring that decisions about AI inclusion or exclusion are aligned with learning intentions, disciplinary epistemologies, regulatory expectations and student development needs. While presented linearly for clarity, the dimensions operate recursively: decisions in one area often necessitate reconsideration of others. The framework therefore supports both programme-level coherence and module-level flexibility, enabling institutions to design assessments that remain meaningful, inclusive and credible in a rapidly evolving technological landscape.

Dimension 1: Purpose and Learning Outcomes in an AI-Mediated Environment

The first dimension emphasises that assessment design must begin with clarity of purpose and alignment with learning outcomes. In an AI-mediated context, this requires educators to articulate which aspects of learning should be supported, enhanced, challenged or safeguarded when students engage with GenAI.

This dimension encourages educators to consider:

- Which learning outcomes genuinely require students to produce work without AI assistance, such as demonstrating foundational disciplinary knowledge, independent reasoning or core technical skills.
- Which outcomes may be supported or enhanced through the transparent and scaffolded use of AI tools—for example, iterative drafting, critical comparison, editing, or data manipulation.
- Which outcomes explicitly require students to evaluate, critique or apply GenAI, reflecting the growing importance of AI literacy as a graduate attribute.

Rather than assuming that GenAI inherently undermines learning, this dimension positions AI as one of many tools that may or may not be appropriate depending on the intended cognitive, ethical and practical demands of a task. By grounding assessment decisions in the purpose of learning, educators can avoid the common pitfall of designing tasks around the technology itself, instead maintaining pedagogic intentionality.

This dimension also recognises that assessment serves purposes beyond the demonstration of predefined competencies. As discussed earlier, assessment may also create opportunities for students to exercise agency, develop distinctive perspectives and engage meaningfully with uncertainty (Biesta, 2009). While constructive alignment remains important, AI mediated assessment environments should avoid reducing learning solely to tightly prescribed outputs that risk narrowing opportunities for intellectual development. Consideration should therefore be given not only to what students must demonstrate, but also to how assessment creates space for judgement, interpretation and individual engagement.

Dimension 2: Authenticity and Professional Relevance

The second dimension focuses on the authenticity of assessment tasks and their alignment with real-world professional practice. The literature shows that many assessment formats common in UK higher education—particularly decontextualised essays and reports—are increasingly vulnerable to AI replication while offering limited opportunities for students to demonstrate judgment, decision-making or application (QAA, 2023). This dimension therefore calls for assessments that reflect the messy, contextual and collaborative nature of contemporary knowledge work.

Examples include:

- Workplace-style tasks requiring students to interpret ambiguous problems, make choices, justify decisions and evaluate outcomes.
- Process-rich assessments that capture planning, iteration and reflection.
- Tasks that integrate multimodal evidence (audio, video, portfolio artefacts) that better mirror real professional practice.
- Assessments that require human-led judgement, interpersonal communication, ethical reasoning or practical application—areas where AI remains limited.

Authenticity in this framework does not imply making assessment “AI-proof”; rather, it involves designing tasks where AI use is integrated into the workflow in ways that reflect responsible professional practice. For disciplines where AI is already influencing industry standards—such as business analytics, education, nursing, software design or media production—AI-inclusive assessments may more accurately prepare students for the expectations of their future roles.

Authenticity in this framework does not imply making assessment “AI proof”; rather, it involves designing tasks where AI use is integrated into the workflow in ways that reflect responsible professional practice. For disciplines where AI is already influencing industry standards, AI inclusive assessments may more accurately prepare students for the expectations of their future roles. Beyond professional relevance, authentic assessment may also create opportunities for students to engage with uncertainty, exercise contextual judgement and develop distinctive responses to complex problems. Preserving space for interpretation, agency and individual engagement may be particularly important within AI mediated environments, ensuring that assessment supports broader educational purposes beyond the demonstration of predefined competencies alone (Biesta, 2009).

Dimension 3: Transparency of AI Use Expectations

The third dimension addresses one of the most critical issues identified in recent UK studies; students and staff frequently do not share a common understanding of what constitutes acceptable AI use. This misalignment creates risk, confusion and inequity (Freeman, 2025; Jisc, 2025). Transparency is therefore essential to ethical and inclusive assessment.

This dimension requires that assessment briefs and marking criteria explicitly state:

- The role AI is permitted to play (e.g., idea generation, editing, coding, data modelling).
- The role AI is prohibited from playing (e.g., writing entire submissions, generating solutions, replacing required disciplinary reasoning).
- Whether AI use is required, and if so, what tools or processes students must engage with.
- How students should document, cite or reflect on their AI use.
- How AI use will be evaluated or assessed.

Transparency also extends to staff; internal guides, exemplars, and programme-level policies are necessary to ensure consistent interpretation across modules. This dimension enables institutions to address the gap between policy intent and student experience, reducing unintentional misconduct and supporting ethical AI literacy.

Transparency also extends beyond written guidance and policy statements. Research suggests that transparency through written criteria or rubrics alone is often insufficient, with students and tutors developing richer understandings of standards through dialogue, exemplars and shared discussion of quality (Rust, Price and O'Donovan, 2003; O'Connell et al., 2016). Bearman and Ajjawi (2021) similarly argue that calibration activities can support the development of evaluative judgement by enabling participants to compare, interpret and discuss examples of assessed work. In AI mediated environments, opportunities to engage critically with exemplars, including AI assisted and AI generated outputs, may become increasingly important in supporting shared understandings of acceptable practice. Transparency therefore extends beyond simply specifying rules around AI use and instead involves developing students' assessment literacy, evaluative judgement and ethical AI capability. This dimension enables institutions to address the gap between policy intent and student experience, reducing unintentional misconduct and supporting ethical AI literacy.

Dimension 4: Assessment Modes and Evidence Types

The fourth dimension emphasises the need to consider alternative assessment modes and evidence types in ways that maintain validity and support richer demonstrations of learning. Given the generative capabilities of AI, reliance on a narrow set of text-based outputs risks both invalidation and inequity. This dimension supports programme teams to broaden their assessment repertoire.

Potential assessment modes include:

- **Annotated portfolios** combining AI-generated and student-generated elements, with commentary on choices and decisions.
- **Critical evaluations** of AI outputs, asking students to compare, critique or correct AI-generated artefacts.
- **Design logs**, where students chronicle their process, decisions and refinements.
- **Recorded oral assessments**, viva-style defences, or multimodal explanations.
- **Collaborative tasks** that integrate AI as one of several tools used in the workflow.

This broadening strengthens validity by capturing aspects of learning that AI cannot reproduce, such as reasoning, dialogue, reflection and adaptation. It also supports students with varying strengths, learning styles and linguistic backgrounds.

This broadening of assessment evidence strengthens validity by capturing aspects of learning that generative AI cannot easily reproduce, including reasoning, dialogue, reflection and adaptation. Recent sector guidance has argued that traditional text-based assessments may no longer provide sufficient evidence of student learning in AI enabled environments and has called for greater emphasis on application, process and judgement (QAA, 2023). Similarly, Kofinas, Tsay and Pike (2025) argue that authenticity alone does not protect assessment integrity and that assessment tasks increasingly require opportunities for contextual decision making and critical engagement. Reconsidering assessment modes may therefore serve not only a technical function in responding to AI capabilities, but also a pedagogic one. Broader forms of evidence can create opportunities for students to demonstrate agency, interpretation and meaningful engagement with learning, supporting richer demonstrations of knowledge and preserving educational purposes beyond standardised performance measures (Biesta, 2009). Such approaches may also align with wider traditions of authentic and sustainable assessment that emphasise process, judgement and application rather than product alone (Sambell, Brown and Graham, 2017; Perkins et al., 2024).

Dimension 5: Integrity, Equity and Accessibility

The fifth dimension reflects the ethical foundation of the framework. It addresses the need for assessment to uphold academic standards while ensuring equity and fair access, aligning with UNESCO, Jisc and QAA principles.

This dimension encourages educators to evaluate:

- Whether all students have equitable access to the AI tools required for an assessment.
- Whether AI-integrated tasks include scaffolding to support students with lower digital capital.
- How detection technologies will be used responsibly, if at all, and how false positives will be mitigated.
- How to design assessments that minimise harm, bias and exclusion.
- How to maintain academic standards without privileging students who enter higher education with greater AI fluency.

This dimension is also where the ethical use of AI in developing the framework is modelled: AI enables efficiency in routine tasks, but human judgement, interpretation and conceptual reasoning remain central. The integration of integrity and equity considerations ensures that AI-enabled assessment supports student success rather than exacerbating inequalities or compromising the meaning of awards.

This dimension also reflects growing concerns that generative AI may reproduce or deepen existing educational inequalities if issues of access, confidence and AI literacy are not explicitly addressed. UNESCO (2023) warns that unequal access to AI tools and digital infrastructure may exacerbate existing inequities, while UK evidence suggests significant differences in students' confidence and engagement with AI technologies (Freeman, 2025; Jisc, 2025). Without careful design, AI integrated assessment risks privileging students with greater existing digital capital and creating new forms of hidden curriculum around acceptable or effective AI use.

The integration of integrity and equity considerations therefore serves not only to protect academic standards but also to support fair participation and meaningful engagement. As discussed earlier, preserving opportunities for agency and evaluative judgement remains important within AI mediated environments (Biesta, 2009). AI may enhance efficiency and support learning processes, but human judgement, interpretation and critical reasoning remain central. Assessment design must therefore

ensure that AI enabled practices support student success rather than reproducing inequities or compromising the meaning and credibility of awards.

Illustrative application of the GenAI Integrated Assessment Design Framework

To demonstrate the practical application of the framework, this section presents an illustrative example showing how the five dimensions may be used to review and redesign an assessment task. The purpose is not to report empirical implementation or evaluate outcomes, but to provide a practical illustration of how educators and programme teams might apply the framework within assessment design processes.

Context

Consider a Level 6 undergraduate Education module in which students are required to produce a traditional 3000-word essay exploring a contemporary issue in inclusive education. The assessment is designed to develop students' critical engagement with policy, theory and practice and aligns with programme learning outcomes relating to analysis, synthesis and evaluation. While the assessment supports academic writing and disciplinary engagement, the emergence of generative AI raises questions regarding whether a traditional essay continues to provide sufficiently rich evidence of student learning. Recent sector guidance has suggested that decontextualised text-based assessments may become increasingly vulnerable to AI supported production and may not always capture the broader processes of reasoning, judgement and application that higher education seeks to develop (QAA, 2023).

Applying the GenAI Integrated Assessment Design Framework may support educators in critically reviewing whether the assessment remains aligned with its intended purpose and whether opportunities exist to strengthen authenticity, transparency and educational value.

Dimension 1. Purpose and Learning Outcomes

The first stage of review involves reconsidering the intended purpose of the assessment and clarifying what forms of learning students are expected to demonstrate. Reflection through this dimension may prompt consideration of whether the existing assessment primarily evidences written production or whether it creates opportunities for students to demonstrate critical judgement, interpretation and independent thinking.

For example, programme teams may identify that the intended outcomes extend beyond academic writing and include professional judgement, contextual application and critical evaluation. This raises questions regarding whether these outcomes are adequately captured through a single essay format. Consideration may therefore be given to whether assessment design creates opportunities not only for the demonstration of competence, but also for student agency and intellectual engagement.

Dimension 2. Authenticity and Professional Relevance

Review through the authenticity dimension may encourage consideration of whether the task reflects the kinds of activities students are likely to encounter within professional contexts. Rather than requiring students to write a decontextualised academic essay, the assessment could be reframed around a professional scenario.

For instance, students might adopt the role of an educational consultant asked to provide evidence informed recommendations for improving inclusion within a school setting. Students may be required to interpret contextual information, evaluate competing priorities and justify decisions in relation to educational theory and practice.

This approach introduces greater ambiguity and contextual judgement while retaining alignment with module learning outcomes. It also creates opportunities for students to engage with authentic professional practices rather than simply reproduce disciplinary knowledge.

Dimension 3. Transparency of AI Use Expectations

Application of the framework also highlights the importance of making expectations regarding AI use explicit. Rather than leaving students to interpret institutional guidance independently, the assessment brief could clearly communicate what forms of AI use are permitted, prohibited or required.

For example, students may be permitted to use generative AI tools for initial idea generation, brainstorming or identifying relevant areas for further reading. However, students might be required to document this use through a short AI use declaration or reflective commentary explaining how AI informed their thinking and how outputs were critically evaluated.

Exemplars and guided discussion activities could also be incorporated to support shared understandings of acceptable practice. Students may engage with examples of AI assisted work and discuss issues relating to quality, judgement and responsible use.

Dimension 4. Assessment Modes and Evidence Types

Reviewing assessment modes may encourage consideration of whether richer forms of evidence could support more comprehensive demonstrations of learning. Rather than relying exclusively on a final written product, additional process focused elements could be incorporated.

For example, alongside the consultant report, students might be asked to submit a short five-minute recorded rationale explaining how recommendations were developed and justified. Students could also include a reflective commentary documenting key decisions made throughout the task, including where AI tools were used, how outputs were evaluated, and why particular ideas were accepted, adapted or rejected. In one scenario, a student may use generative AI to produce initial suggestions for improving inclusive practice, then critically evaluate these suggestions against educational theory and the specific school context before explaining why some recommendations were modified or discarded.

Opportunities could also be created for students to critique AI generated outputs directly, identifying limitations, inaccuracies or contextual issues that require human judgement. Broadening evidence in this way may support stronger demonstrations of reasoning, interpretation and evaluative judgement while reducing overreliance on polished written outputs alone.

Dimension 5. Integrity, Equity and Accessibility

The final dimension prompts consideration of how redesign may affect different student groups. AI integrated assessment cannot assume equal access to technologies, confidence or prior experience. Review through this dimension may therefore encourage programme teams to provide scaffolded guidance around AI use, opportunities for skill development and explicit discussion of ethical considerations. Students with limited familiarity with AI tools may require additional support to ensure that expectations do not unintentionally privilege those entering higher education with greater digital capital.

Equity considerations may also extend beyond access to technology and include ensuring that redesigned assessments remain inclusive and accessible for diverse learners.

Summary and Reflective Considerations

Viewed collectively, the framework shifts the focus from attempting to make assessments resistant to AI towards more fundamental questions regarding assessment purpose and evidence. Rather than positioning AI solely as a threat to academic integrity, the framework encourages educators to reconsider how assessments create opportunities for judgement, interpretation and meaningful engagement with learning. The illustrative example presented here is not intended as evaluation but rather demonstrates how the framework may support practical assessment redesign by connecting wider theoretical debates around authenticity, agency and equity with everyday assessment decision making.

Together, these five dimensions form a coherent and adaptable model that bridges assessment theory, AI ethics, policy expectations and practical classroom realities. The framework positions GenAI as a catalyst for reflective and principled assessment design, supporting universities to move beyond reactive and risk oriented responses towards approaches grounded in authenticity, transparency, equity and meaningful student engagement. In doing so, the framework seeks not only to respond to technological disruption but also to reconnect assessment with broader educational purposes, including agency, judgement and student development within increasingly AI mediated learning environments.

Discussion

The GenAI-Integrated Assessment Design Framework responds to a rapidly evolving UK higher education landscape in which generative AI is becoming embedded within student learning practices, institutional policy, and quality assurance discourse. UK sector bodies such as the QAA (2023), Jisc (2023) and OfS (2025) have consistently emphasised that existing assessment strategies require urgent reconsideration to remain valid, fair and credible in the context of generative AI. This framework offers a structured, pedagogically grounded response to these calls, translating high-level guidance into actionable principles for assessment design.

Central to this contribution is the shift from risk-dominated narratives towards a more developmental understanding of assessment in an AI-enabled environment. UK literature has highlighted how early institutional responses frequently prioritised detection technologies and restrictive policies, often reinforcing a culture of surveillance rather than learning (QAA, 2023; Jisc, 2024). Such approaches risk undermining trust and may disproportionately disadvantage students who already experience educational marginalisation. The proposed framework reframes integrity as an outcome of thoughtful design, transparency and scaffolded skill development, aligning with emerging UK scholarship that advocates for educative rather than punitive models of academic integrity (Bennett and Abusalem, 2024).

The framework also aligns with recent UK evidence indicating that students are already incorporating AI into their assessment workflows, often in ways that are pragmatic but ethically ambiguous. The HEPI Student AI Surveys demonstrate that AI use amongst UK undergraduates is widespread and accelerating, yet many students remain uncertain about what constitutes acceptable practice (Freeman, 2024; 2025). Jisc's qualitative studies further highlight this uncertainty, identifying inconsistent staff messaging and unclear assessment guidance as key sources of anxiety for students (Jisc, 2024). By prioritising transparency as a core design dimension, the framework directly addresses this gap, supporting clearer communication of expectations around AI use and reducing the likelihood of inadvertent misconduct.

The framework also engages meaningfully with UK debates around authentic assessment. Although authentic assessment has long been positioned as a solution to academic dishonesty (Sambell, Brown and Graham, 2017), recent UK discourse suggests that authenticity alone does not inoculate against

AI misuse. QAA (2023) cautions that many so-called “authentic” tasks remain vulnerable if they continue to rely primarily on text-based outputs without requiring critical judgement, process evidence or personal contextualisation. The framework therefore expands orthodox interpretations of authenticity by embedding reflective engagement with AI, decision-making processes and critical evaluation as core features of assessment design, ensuring that tasks remain pedagogically robust even where AI is present.

Equity is another area where the framework offers value within the UK context. Evidence from HEPI and Jisc illustrates that access to and confidence in AI tools is uneven across student populations, with disparities linked to socio-economic background, gender and prior educational experience (Freeman, 2025; Jisc, 2025). Without deliberate design intervention, AI-integrated assessment risks further entrenching inequality. By embedding equity and accessibility as a foundational dimension rather than an add-on, the framework supports assessment practices that actively mitigate digital divides and promote inclusive participation.

From a quality assurance perspective, the framework offers practical potential for integration into established UK processes such as programme approval, course monitoring and enhancement, annual academic review and assessment mapping. Its structured dimensions provide a lens through which programme teams can evaluate both individual assessments and assessment portfolios holistically, ensuring coherence with regulatory expectations and institutional strategy. This aligns with OfS priorities around maintaining the credibility of awards and safeguarding the academic standards underpinning student outcomes (OfS, 2025).

Nevertheless, the framework also acknowledges its limitations. As a conceptual contribution, it does not claim universal applicability or immediate empirical validation. Rather, it functions as a flexible heuristic, designed to evolve alongside technological developments and pedagogic practice. UK institutions differ widely in resources, digital infrastructure, staff confidence and disciplinary context, all of which will shape how the framework is interpreted. This reinforces the importance of ongoing institutional dialogue, reflective practice and iterative refinement.

Importantly, the framework positions generative AI not as a binary threat to traditional assessment, but as a catalyst for renewed critical engagement with the purposes of assessment itself. By foregrounding learning outcomes, authenticity, transparency and equity, it supports a shift towards assessment practices that better reflect contemporary knowledge work and graduate realities in the UK context. In doing so, it contributes to a growing scholarly and policy consensus that the future of assessment lies not in resisting technological change, but in shaping it responsibly, critically and ethically.

Implications for Practice and Policy

The GenAI-Integrated Assessment Design Framework has significant implications for assessment practice, institutional policy and quality enhancement across UK higher education. As generative AI becomes increasingly embedded in student learning behaviours, the framework offers a structured means through which higher education providers can move from reactive, risk-oriented strategies towards coherent, values-led approaches to assessment redesign.

Implications for pedagogic practice

At the level of teaching and learning, the framework encourages educators to engage in more intentional, reflective assessment design. Rather than attempting to make assessments “AI-proof”, staff are invited to design tasks that explicitly incorporate or respond to AI in ways that develop higher-order thinking, critical judgement and ethical awareness. This aligns with Jisc’s assertion that

universities should prioritise “AI-ready pedagogy” which develops students’ capacity to work critically alongside AI tools, rather than seeking to isolate them from such technologies (Jisc, 2023).

The framework also provides a practical mechanism for strengthening assessment literacy among staff. UK studies highlight that academic confidence in using AI for assessment design varies significantly, with many staff expressing uncertainty about how to adapt their practice (QAA, 2023; Jisc, 2024). By offering a clear set of design dimensions, the framework can act as both a developmental tool and a reflective prompt, enabling staff to interrogate the purpose, authenticity and equity of their assessment strategies. This supports a shift from ad hoc experimentation towards more coherent pedagogic practice.

Importantly, the framework positions transparency as central to effective learning. Clear communication of AI expectations within assessment briefs, marking criteria and student guidance supports more equitable learning environments and reduces the risk of inadvertent misconduct. This resonates with HEPI findings that students frequently desire clearer institutional guidance on acceptable AI use and feel more confident where expectations are explicit (Freeman, 2024; 2025).

Implications for programme design and curriculum coherence

At programme level, the framework provides a valuable lens through which assessment portfolios can be reviewed holistically. UK sector guidance increasingly stresses the importance of assessment mapping and coherence, particularly in relation to maintaining standards, reducing assessment burden and supporting progression (QAA, 2023). The framework offers a way to evaluate how AI-integrated assessments align with programme learning outcomes, ensuring that AI use is scaffolded progressively and coherently across levels of study.

This has implications for curriculum design, particularly in disciplines where AI is rapidly reshaping professional practice, such as education, health, business and creative industries. The framework supports curriculum teams to embed AI literacy as a developmental thread rather than an isolated skill, aligning assessment design with employability outcomes and sector expectations. In this sense, the model supports graduate preparedness while safeguarding academic integrity and disciplinary standards.

Implications for institutional policy and governance

At an institutional level, the framework offers a principled foundation for policy development. While many UK institutions have developed AI policies, these frequently remain high-level and lack clear operational guidance for assessment design (QAA, 2023). The framework can inform the creation of:

- Assessment design guidance tailored for the AI era
- Policy statements on permissible AI use in assessment
- Staff training resources and toolkits
- AI clauses within standardised assessment brief templates

By aligning assessment design policy with pedagogic values rather than purely regulatory concerns, institutions can create more coherent and ethically grounded approaches. This also supports compliance with OfS expectations regarding quality, standards and the credibility of awards (OfS, 2025), ensuring that qualifications remain meaningful and defensible in an AI-enabled context.

Implications for quality assurance and enhancement

The framework has relevance for quality assurance and enhancement processes within UK higher education. It can be integrated into:

- Programme approval and re-approval processes
- Course monitoring and enhancement
- Periodic review and validation panels
- Assessment mapping exercises
- External examiner considerations

By providing a structured lens through which assessment portfolios can be evaluated, the framework supports evidence-based dialogue around the appropriateness, robustness and equity of assessment strategies. This aligns with QAA's growing emphasis on sustainable assessment practices and continuous enhancement rather than static compliance.

Furthermore, the framework enables institutions to demonstrate proactive engagement with AI-related risks, supporting defensible responses to external scrutiny from regulators and professional bodies. It thus contributes to a culture of quality that is dynamic, reflective and responsive to technological change.

Implications for equity and student experience

From a student perspective, the framework has the potential to enhance clarity, fairness and confidence in assessment processes. By embedding explicit expectations around AI use and designing tasks that are accessible and inclusive, institutions can reduce ambiguity and anxiety surrounding AI engagement. This is particularly important given evidence that students from disadvantaged backgrounds often report lower confidence and higher anxiety around AI use in assessment contexts (Freeman, 2025; Jisc, 2025).

The framework also supports more consistent student experience across modules and programmes, addressing concerns raised by UK students regarding inconsistent messaging and uneven expectations across courses (Jisc, 2024). By promoting transparency and coherence, the model supports greater trust in assessment processes and strengthens students' understanding of academic integrity.

Implications for future research and sector development

While the framework offers a theoretically informed contribution, it also opens up significant opportunities for future research. Empirical studies could explore its implementation across different disciplines, evaluate its impact on student outcomes, and examine staff experiences of assessment redesign using the model. Collaborative, sector-wide studies could further refine the framework and explore its applicability across diverse institutional contexts.

Conclusion

The rapid integration of generative artificial intelligence into higher education represents one of the most profound disruptions to assessment practice in recent decades. As UK students increasingly incorporate AI tools into their academic work, and institutions grapple with the implications for standards, integrity and equity, the need for coherent, pedagogically grounded approaches to assessment design has become urgent. This paper has argued that existing responses — often characterised by attempts to regulate, detect or prohibit AI use — are insufficient to address the deeper structural and educational challenges posed by GenAI.

In response, the paper has introduced the GenAI-Integrated Assessment Design Framework, a conceptual model designed to support higher education providers in navigating the complexities of assessment in an AI-enabled environment. The framework offers a structured approach centred on five interdependent dimensions: purpose and learning outcomes, authenticity and professional relevance, transparency of AI use expectations, assessment modes and evidence types, and integrity, equity and accessibility. Together, these dimensions provide a coherent lens through which assessment

can be reimagined not as a site of technological threat, but as a space for intentional, ethical and future-focused pedagogic redesign.

The framework makes several original contributions; conceptually, it bridges established assessment theory with emerging scholarship on generative AI and aligns with evolving UK policy guidance from QAA, Jisc and the Office for Students. Practically, it offers a usable tool for educators, programme leaders and quality professionals seeking to redesign assessment portfolios in ways that preserve academic standards while recognising the realities of AI-mediated learning. By embedding equity and transparency at its core, the model also responds to documented disparities in AI access and literacy, supporting more inclusive and just assessment practices.

At a broader level, the framework reinforces the imperative for higher education to shift from reactive and defensive stances towards more reflective and principled engagement with AI. Rather than attempting to restore an assessment landscape that no longer exists, institutions are encouraged to embrace complexity and re-centre assessment around human judgement, critical engagement and ethical reasoning. In doing so, assessment becomes not merely a measure of knowledge, but a vehicle for developing the very competencies that graduates will require to navigate AI-rich professional and societal environments.

While this paper provides a theoretically grounded starting point, further empirical research is needed to explore how the framework functions across different disciplines, institutional contexts and student populations. Longitudinal studies examining its impact on learning outcomes, student confidence and academic integrity would offer valuable insights for refinement. Collaborative sector-wide engagement could further evolve the model and support the development of shared best practice.

In conclusion, the GenAI-Integrated Assessment Design Framework offers a timely and principled contribution to the evolving discourse on assessment in higher education. By positioning generative AI as both a challenge and an opportunity, the framework supports universities in crafting assessment practices that are robust, equitable and aligned with the demands of an AI-enabled world — ensuring that assessment remains meaningful, credible and educationally transformative.

References

- Adams, R. (2024) Researchers fool university markers with AI-generated exam papers. *The Guardian*. 24 June 2024. <https://www.theguardian.com/education/article/2024/jun/26/researchers-fool-university-markers-with-ai-generated-exam-papers> (Accessed: 21 April 2025).
- Attewell, S. (2024) Generative AI for College and University Governors. [Blog post] 17 September. Available at: <https://nationalcentreforai.jiscinvolve.org/wp/2024/09/17/generative-ai-for-college-and-university-governors/> (Accessed: 21 April 2025).
- Bearman, M. and Ajjawi, R. (2021) Can a rubric do more than be transparent? Invitation as a new metaphor for assessment criteria. *Studies in Higher Education*, 46(2), pp. 359–368.
- Bennett, L. and Abusalem, A. (2024) 'Building academic integrity and capacity in digital assessment in higher education', *Athens Journal of Education*, 11(1), pp. 71–94.
- Biesta, G. (2009) 'Good education in an age of measurement: On the need to reconnect with the question of purpose in education', *Educational Assessment, Evaluation and Accountability*, 21(1), pp. 33–46.
- Bittle, K. and El-Gayar, O. (2025) 'Generative AI and Academic Integrity in Higher Education: A Systematic Review and Research Agenda', *Information*, 16 (4), pp. 296.
- Canterbury Christ Church University (2025) *Artificial Intelligence Assessment Scale (AIAS)*. Available at: <https://www.canterbury.ac.uk/learning-and-teaching-enhancement/resources-for-academics/Artificial-Intelligence-Assessment-Scale-AIAS.aspx> (Accessed: 21 October 2025).

- Francis, N.J., Jones, S. and Smith, D.P. (2025) 'Generative AI in Higher Education: Balancing Innovation and Integrity', *British Journal of Biomedical Science*, 81, Article 14048.
- Freeman, J. (2024) *Provide or punish? Students' views on generative AI in higher education (HEPI Policy Note 51)*. Oxford: Higher Education Policy Institute. Available at: <https://www.hepi.ac.uk/wp-content/uploads/2024/01/HEPI-Policy-Note-51.pdf> (Accessed: 16 July 2025).
- Freeman, J. (2025) *Student Generative AI Survey 2025*. London: Higher Education Policy Institute (HEPI).
- Foster, P. and Borrett, A. (2025) 'UK universities warned to stress-test assessments as AI use surges among students', *Financial Times*, 26 February. Available at: <https://www.ft.com/content/d591fb1a-9f6c-4345-b5fc-781e091ae3f8> (Accessed: 01 November 2025).
- Furze, L., Perkins, M., Roe, J. and MacVaugh, J. (2024) 'The AI Assessment Scale (AIAS) in action: A pilot implementation of GenAI supported assessment- A Preprint', *arXiv preprint*, Available at: <https://arxiv.org/abs/2403.14692>? (Accessed: 01 September 2025).
- Gonsalves, C. (2025) 'Contextual assessment design in the age of generative AI', *Journal of Learning Development in Higher Education*, Issue 34, February.
- Halliday, S., Lavis, T., Callaghan, P. and Chur-Hansen, A. (2025) 'Generative AI in higher education psychology programmes: a scoping review exploring the opportunities for its use in assessment methods', *Australian Psychologist*. Advance online publication.
- Jin, Y., Yan, L., Echeverria, V., Gašević, D. & Martinez-Maldonado, R. (2025) 'Generative AI in higher education: A global perspective of institutional adoption policies and guidelines', *Computers and Education: Artificial Intelligence*, 8, 100348.
- Jisc (2023) *Artificial Intelligence in Tertiary Education: A practical guide*. Bristol: Jisc. Available at: <https://www.jisc.ac.uk/guides/artificial-intelligence> (Accessed: 08 November 2025).
- Jisc (2025) *Student Perceptions of Artificial Intelligence in Further and Higher Education*. Bristol: Jisc. Available at: <https://www.jisc.ac.uk/reports/student-perceptions-of-ai> (Accessed: 08 November 2025).
- Kofinas, A. K., Tsay, C. H.-H. and Pike, D. (2025) 'The impact of generative AI on academic integrity of authentic assessments within a higher education context', *British Journal of Educational Technology*, 56 (6), pp. 2522-2549.
- Luo, Y. (2024) 'A critical review of generative AI policies in higher education assessment', *Assessment & Evaluation in Higher Education*, 49(3), pp. 401–416.
- Oc, Y., Gonsalves, C. and Quamina, L.T. (2025) 'Generative AI in higher education assessments: examining risk and tech savviness on student's adoption', *Journal of Marketing Education*, 47(2), pp. 138–155.
- O'Connell, B., De Lange, P., Freeman, M., Hancock, P., Abraham, A., Howieson, B. and Watty, K. (2016). Does calibration reduce variability in the assessment of accounting learning outcomes? *Assessment & Evaluation in Higher Education*, 41(3), pp. 331–349.
- Office for Students (OfS) (2025) *Embracing innovation in higher education: Our approach to artificial intelligence*. London: OfS. Available at: <https://www.officeforstudents.org.uk/news-blog-and-events/blog/embracing-innovation-in-higher-education-our-approach-to-artificial-intelligence/> (Accessed: 09 November 2025).
- Perkins, M., Furze, L., Roe, J. and MacVaugh, J. (2024) 'The Artificial Intelligence Assessment Scale (AIAS): A framework for integrating AI into assessment design', *Assessment & Evaluation in Higher Education*, 49(1), pp. 1–15.
- QAA (2023) *Reconsidering assessment for the ChatGPT era*. Gloucester: Quality Assurance Agency. Available at: <https://www.qaa.ac.uk/docs/qaa/guidance/reconsidering-assessment-for-the-chatgpt-era.pdf> (Accessed: 07 November 2025).
- Quality Assurance Agency for Higher Education (QAA). (2025) *Transnational perspectives on ethical uses of Generative AI (GenAI) in assessment*. Available at:

- <https://www.qaa.ac.uk/membership/benefits-of-qaa-membership/collaborative-enhancement-projects/generative-ai/transnational-perspectives-on-ethical-uses-of-gen-ai-in-assessment> (Accessed: 21 September 2025).
- The Guardian (2023) 'AI essays pass university exams undetected, study finds'. Available at: <https://www.theguardian.com/education/2023/jan/12/ai-chatgpt-university-assessments> (Accessed: 11 November 2025).
- Rust, C., Price, M. and O'Donovan, B. (2003) 'Improving students' learning by developing their understanding of assessment criteria and processes', *Assessment and Evaluation in Higher Education*, 28(2), pp. 147–164.
- Sallai, G., Smith, J. and Crook, C. (2024) 'Generative AI and the future of assessment in UK higher education: Integrity, cognition and pedagogy', *British Journal of Educational Technology*, 55(2), pp. 412–428.
- Sambell, K., Brown, S. and Graham, L. (2017) *Professionalism in practice: Key directions for assessment for learning*. London: Routledge.
- UNESCO (2023) *Guidance on Generative AI in Education and Research*. Paris: UNESCO. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000386693> (Accessed: 06 November 2025).
- University of Warwick (2025) "*The AI Assessment Scale (AIAS) in action: A pilot implementation of GenAI supported assessment*". Available at: https://warwick.ac.uk/fac/cross_fac/eduport/edufund/projects/yang/projects/the-ai-assessment-scale-aias-in-action-a-pilot-implementation-of-genai-supported-assessment-a-b3d58c (Accessed: 21 September 2025).
- Xia, Q., Weng, X., Ouyang, F., Lin, T.J. and Chiu, T.K.F. (2024) 'A scoping review on how generative artificial intelligence transforms assessment in higher education', *International Journal of Educational Technology in Higher Education*, 21, Article 40