



RESEARCH ARTICLE OPEN ACCESS

A Conceptual Framework for Authentic AI-Enabled Assessment (AAA) in Higher Education

Dilani Gedera

Faculty of Business Economics and Law, Auckland University of Technology (AUT), Auckland, New Zealand

Correspondence: Dilani Gedera (dilani.gedera@aut.ac.nz)**Received:** 4 July 2025 | **Revised:** 23 March 2026 | **Accepted:** 9 July 2026**Keywords:** AI-enhanced assessments | authentic assessment | business education | generative AI | higher education

ABSTRACT

Higher education is at a critical juncture as generative artificial intelligence (AI) redefines assessment practices. This article offers a conceptual framework illustrating how authentic AI-enhanced assessments (AAA) can support the reclamation of higher education's core mission: developing critical thinking, creativity, ethical reasoning, and meaningful real-world problem-solving. With the advent of AI, current assessment practices may not adequately evaluate the deep interdisciplinary competencies required for today's complex professional environments. Instead of resorting to increased surveillance or outright bans on AI, this article contends that institutions should embrace AI as a supportive tool to reinforce authentic assessment strategies. The AAA framework emphasises not only the final output but also the iterative learning process—integrating reflective practice, student agency, and responsiveness to unexpected twists that compel students to think on their feet. The article demonstrates that AI-enhanced authentic assessments—ranging from interactive tasks to multi-stage applied projects—can transform evaluation across disciplines by supporting the development of a solid theoretical foundation and the practical, adaptable competencies demanded by modern professional practice.

1 | Introduction

Higher education is at a crossroads. The rapid rise of generative artificial intelligence (AI) tools has sparked both alarm and reflection across higher education institutions. On the one hand, educators worry that these tools enable students to cheat on assessments with unprecedented ease, undermining academic integrity. On the other hand, the advent of powerful AI has prompted a deeper question: What is the true purpose of higher education, and are our current assessment practices aligned with that purpose? For decades, universities have professed a mission beyond the mere transmission of facts—a mission to cultivate critical thinking, creativity, ethical reasoning, and real-world problem-solving skills in their graduates. However, many widely used assessment formats remain poorly aligned with these aims. AI has brought this misalignment into sharper

focus, prompting renewed scrutiny of what assessments are actually asking students to demonstrate. In this sense, AI can be understood not only as a disruption but also as a catalyst for rethinking how higher education assesses learning in ways that remain true to its broader educational purpose.

This article argues that the proper response to AI's encroachment is not to double down on surveillance or ban the use of AI but to guide and structure its responsible use through authentic assessment strategies. Rather than viewing AI purely as a threat to academic integrity, AI can be utilised as a tool to support higher-order learning. Redesigning assessment in this way enables educators to align evaluation more closely with the kinds of judgement, application, and reasoning that higher education claims to value.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Future in Educational Research* published by John Wiley & Sons Australia, Ltd on behalf of Southwest University.

Existing frameworks on authentic assessment provide valuable principles for designing real-world higher-order tasks, whereas emerging AI-related assessment guidance has focused on risk mitigation, academic integrity, and decisions about when AI should be restricted or permitted. What is still missing is an integrated framework that brings these conversations together. This article contributes a new model of AI-enabled authenticity by integrating human–AI collaboration, staged AI integration, and controlled and noncontrolled assessment pathways to support assurance of learning while promoting the ethical and responsible use of AI. In this article, controlled assessments refer to tasks where learning can be directly observed under defined conditions, such as invigilated or supervised settings, whereas noncontrolled assessments involve engagement with broader resources and, where appropriate, AI tools, while still being required to demonstrate achievement of the intended learning outcomes.

Although business education provides the illustrative context in this article, the framework is intentionally designed to be applicable across disciplines and adaptable to diverse fields within higher education.

2 | Literature Review

2.1 | The Core Mission of Higher Education: Beyond Content to Critical Competencies

Universities have long been entrusted with a mission that goes far beyond vocational training or the mere imparting of factual knowledge. Classic and contemporary scholars alike assert that higher education's fundamental purpose is to develop intellectual and moral capacities—to form critical, creative, and socially responsible thinkers. For example, Barnett (1997) argues that the underlying purpose of higher education is to cultivate critical being, enabling students to develop the capacities to think critically, to understand themselves critically and to engage with the world in a critical manner. In a similar vein, philosopher Nussbaum (2010) contends that universities should produce not just technically skilled workers but competent citizens who can approach complex societal issues with creativity and ethical reasoning.

In practice, attributes such as independent thinking, adaptability, and ethical judgement are embedded in many universities' graduate capability frameworks and accreditation standards (e.g., AACSB for business schools). Business education, in particular, faces calls to ensure graduates can navigate ambiguous problems, exercise ethical leadership, and innovate in dynamic markets. The knowledge economy and global society demand graduates who can not only recall knowledge but also apply it in novel situations, critique information and continuously learn. Thus, the core mission of higher education centres on preparing students to engage with complex, uncertain and evolving real-world contexts. Achieving this mission requires alignment between curriculum, pedagogy and assessment.

2.2 | Limitations of Traditional Assessment in Fulfilling the Purpose of Higher Education

Assessment plays an important role in shaping student learning. As Boud and Associates (2010) noted, assessment has a

powerful influence on what and how students learn; it is often said that “assessment drives learning”. If our assessments predominantly reward rote memorisation or formulaic responses, we should not be surprised when students focus on lower-order learning despite espoused goals of critical thinking. Traditional forms of assessment commonly used in higher education—such as closed-book exams, essay assignments, and multiple-choice question (MCQ) tests—have been widely critiqued for their misalignment with higher-order learning outcomes (Black and Wiliam 2009; French et al. 2024; Gedera 2023). Assessment formats, which typically emphasise factual recall and lower-order cognitive skills, are increasingly vulnerable to AI-generated content (Evangelista 2025).

Scholars argue that traditional assessment formats often encourage surface learning and the reproduction of information rather than its application, thereby limiting the development of transferable understanding (French et al. 2024). They are frequently detached from real-world contexts, meaning students are not required to demonstrate how knowledge is used in practice (Gulikers et al. 2004; Villarroel et al. 2018), contributing to a recognised gap between theoretical knowledge and the ability to address real-world problems (Gonzales 2020; Singh et al. 2014). These approaches are typically summative, offering limited opportunities for formative feedback, reflection or the development of metacognitive skills essential for lifelong learning (Gedera 2023). In addition, standardised formats often constrain student agency, reducing opportunities for autonomy, engagement and the demonstration of individual strengths (Black and Wiliam 2009). More recently, technological developments have further exposed these limitations because such tasks can be completed without clear evidence of students' own understanding or reasoning (Evangelista 2025).

In business education, these issues are particularly pronounced. Business programmes historically relied on case-study essays and exams that may not capture the messy complex interdisciplinary nature of real business challenges. For example, a multiple-choice test in accounting might assess technical concepts, but in practice, an accounts manager must integrate knowledge, communicate with nonaccounting colleagues and make ethical decisions under uncertainty—aspects not captured by traditional tests.

There is a clear misalignment between traditional assessment methods and the core mission of higher education. Addressing this requires assessment practices that explicitly elicit and recognise the kinds of judgement, application, and understanding that universities seek to develop. This realisation has led to growing calls for assessment reform. Even prior to the AI revolution, higher education institutions were shifting towards more authentic student-centred assessment models (Biggs and Tang 2011; Boud and Associates 2010; Gedera 2023; Wiggins 1990). The next section explores authentic assessment as a leading paradigm for aligning assessment with these educational goals.

2.3 | Authentic Assessment: Definition and Key Characteristics

Authentic assessment emerged in the late 20th century as a response to the shortcomings of traditional assessments. Grant

Wiggins (1990) was an early advocate, arguing that assessments should engage students in “worthy tasks” that mirror the challenges faced in the real world, rather than decontextualised drills. Authentic assessment refers to the design of assessment tasks that resemble real-life applications of knowledge and skills, requiring students to construct responses and perform or produce work as they would in a professional or societal context. Mueller (2005) succinctly described authentic assessment as asking students to “do the subject”—for example, instead of recalling accounting rules, actually analyse financial statements; instead of writing about marketing, actually develop a marketing plan for a product.

Over the years, scholars have elaborated on the defining characteristics of authentic assessment (Ashford-Rowe et al. 2014; Gedera 2023; Gulikers et al. 2004; Villarroel et al. 2018). Key characteristics include:

- **Real-world relevance:** The task is embedded in a context that simulates real-life complexity. It addresses problems or challenges that professionals or citizens genuinely face outside the classroom. For instance, students might be asked to consult for a local business, analyse a case based on current events, or role-play a decision-making scenario in corporate governance. This relevance increases student motivation and helps them see the purpose of their learning.
- **Higher-order thinking:** Authentic tasks demand analysis, synthesis, evaluation, and creativity. They typically have no single correct answer, requiring students to formulate arguments or solutions and justify their approach. This stands in contrast to selecting a choice in an MCQ. The need to transfer learning to a new context demonstrates that knowledge transfer has occurred (a high-level outcome in Bloom's taxonomy). For example, a student may need to apply economic theory to a new market scenario or use psychological concepts to design a marketing campaign.
- **Performance or product- and process-based:** Rather than solely writing an exam or essay, authentic assessment requires students to produce a tangible product—such as a presentation, portfolio or prototype—or perform a task such as delivering a pitch or conducting a role-play negotiation. The outcome is often a polished piece of work that holds value beyond the classroom, aligning with professional practice where competence is demonstrated through concrete outputs. Equally important, authentic assessments evaluate the process involved in producing the final work. This means assessing the iterative development, strategic planning, and reflective practices that students engage in as they work toward their final outcome. By analysing both the product and the process, educators gain insight into how students develop critical thinking, problem-solving and metacognitive skills that are essential for real-world professional performance.
- **Open-ended and complex problems:** Authentic assessments often involve complexity and ambiguity. They might be ill-structured problems with multiple viable approaches, requiring students to define the problem and scope themselves. This mimics real situations where problems are not neatly packaged.

- **Use of real tools and resources:** Students are usually allowed (even encouraged) to use the same tools, resources and references that they would use in the real world. This could include calculators, software, databases and—notably in today's context—AI tools, as long as their use is transparent, ethical and appropriate. By allowing such resources, authentic tasks shift focus from recall to the effective use of information, reflecting modern workplaces where employees have information at their fingertips but must deploy it critically. As an example, an authentic assessment in a strategy course might permit students to use business intelligence databases or AI analysis to inform their strategic plan because a real strategist would do the same. However, an added step would be that students critically evaluate and refine any AI-generated analysis and apply their own professional judgement, ensuring the output reflects AI literacy, application and implications rather than uncritical reliance.
- **Student agency and collaboration:** Authentic tasks can incorporate student choice (e.g., choosing a project topic or the medium of output) and often involve collaboration or interaction. Working in teams, consulting with stakeholders or engaging in peer review are common elements, mirroring professional collaboration. Although collaboration must be structured to allow individual accountability, it adds realism—in business settings, teamwork and communication are vital skills. Moreover, discussing and receiving feedback during the task is natural in professional work, and assessments can include mechanisms for feedback and iteration.
- **Criteria that reflect expertise:** The evaluation of authentic tasks is based on criteria that experts or employers would value. Rubrics might include indicators such as effective communication with a client, depth of analysis, creativity of solution and ethical reasoning—criteria aligned with quality work in the field. Most importantly, an additional criterion is the ability to use AI ethically and appropriately. Gulikers et al. (2004) emphasise the importance of aligning the assessment standards with real-world performance standards (e.g., using industry benchmarks or having an external client judge part of the work for authenticity).
- **Reflection and metacognition:** Many authentic assessments incorporate a reflective element, asking students to articulate what they learned, how they approached the problem and how they evaluated their own performance. This deliberate reflection not only develops metacognitive skills—enabling students to think about their thinking—but also ensures they derive substantive lessons from the experience beyond simply completing the task. For instance, an interactive oral discussion with the student can be used to deepen this reflective process. In such a discussion, an educator engages the students in a dialogue that probes the rationale behind their decisions, challenges them to explain the evolution of their thinking and prompts them to evaluate alternative strategies they might have employed. This dynamic real-time exchange provides immediate feedback, encourages deeper introspection and reinforces the student's ability to articulate and refine their learning process in a manner that mirrors authentic professional interactions.

These characteristics position authentic assessment as a means of aligning assessment with higher-order learning and real-world application, consistent with broader calls for the development of transferable skills such as critical thinking, problem-solving and adaptability (World Economic Forum 2020). Empirical research studies further suggest that such approaches can enhance student engagement and support the development of relevant capabilities (Gulikers et al. 2004; Vlachopoulos and Makri 2024).

However, implementing authentic assessment is not without challenges. The literature points to barriers such as staff workload for designing and grading complex tasks, the need for professional development to create effective rubrics and sometimes resistance from staff and accreditation bodies accustomed to traditional exams (Gedera 2023). Additionally, ensuring consistency and fairness in grading authentic work requires clear rubrics and possibly moderation because student outputs may be highly varied. Despite these challenges, the momentum toward authentic assessment has grown, especially in professional fields like business, engineering, and health, where stakeholders demand proof of practical skills. Recent shifts in higher education delivery, including the expansion of online learning, have further accelerated this trend.

2.4 | AI Disruption in Assessment: Challenge and Opportunity

The release of OpenAI's ChatGPT in late 2022 sent shockwaves through academia. Suddenly, an average student could input an essay prompt and receive a coherent well-structured response in seconds. Early analyses found that ChatGPT could produce passable answers to many typical university assignments, from short essays to coding tasks, often evading plagiarism detectors by generating original text. This raised immediate concerns about a new wave of academic misconduct facilitated by AI. Reports emerged of students submitting AI-written essays, leading universities to scramble for detection tools and policy responses (Cotton et al. 2024).

The initial reactions to AI in education ranged from panic to pragmatism. Some institutions and educators considered strict bans on AI usage for assignments or a wholesale return to in-person invigilated exams and oral assessments. Scholars such as Rudolph et al. (2023) argue that a policing approach alone is inadequate and arguably regressive. Simply banning AI or increasing surveillance treats the symptom (cheating), not the cause (misaligned assessment) and might drive misconduct further underground. Over-reliance on closed-book exams goes against the trend of authentic learning and does not reflect how knowledge and skills are used in professional contexts.

Early responses in higher education often assumed that AI would be most effective on structured, generic or lower-order tasks, whereas assessments requiring critical insight, ethical reasoning or original application would remain more resistant (Dwivedi et al. 2021; Cotton et al. 2024). More recent AI developments, however, have expanded these capabilities, with newer models showing stronger reasoning, long-context analysis, multimodal processing and increasingly agentic performance on complex professional tasks.

Academic and regulatory bodies have begun issuing guidance that reflects these insights. TEQSA's 2023 guidance and 2024 follow-up emphasise that simply trying to "outrun" or ban AI is futile; instead, institutions should adapt by designing assessments that either control AI use or embrace it in a pedagogically sound way (J. M. Lodge 2024; J. Lodge et al. 2023). The University of Sydney's "two-lane approach" is one notable framework, categorising assessments into Lane 1 (controlled assessments for testing foundational knowledge or individual skills under defined conditions) and Lane 2 (noncontrolled assessments that integrate AI use, focusing on higher-order skills and the use of AI). Such approaches highlight the importance of balancing assessment conditions to both assure learning outcomes and support the development of capabilities relevant to AI-rich professional contexts.

AI's capacity to foster critical thinking, support adaptive learning and deliver tailored feedback can enhance assessment practices to more closely mirror real-world inquiry and engagement (Ishmuradova et al. 2025). Importantly, treating AI as a collaborative tool rather than a threat reinforces the development of AI literacy as a key graduate attribute. AI literacy is defined as the ability to understand and use AI tools responsibly, including recognising their strengths, limitations and ethical implications. In business education, where graduates increasingly rely on AI for data analysis, report writing, decision support and other tasks, cultivating AI literacy is paramount. Forward-thinking educators argue that assessments should actively teach and evaluate AI literacy—not merely protect against AI misuse (Holmes and Tuomi 2022; Perrotta and Selwyn 2020).

3 | Conceptual Framework

3.1 | Authentic Assessment as the Nexus of Higher Education's Purpose and AI Integration

Drawing on the literature and context above, this article proposes a conceptual framework—Authentic AI-Enabled Assessment (AAA), which brings together higher-order learning outcomes, authentic assessment strategies and AI integration into a coherent structure for designing authentic assessments in AI-rich learning environments (see Figure 1). In this model, authentic assessment is not an end in itself but functions as the connection point between higher education's purpose and the effective use of AI, emphasising the development of graduates who can think critically, act ethically and work confidently with AI-enabled tools. At the heart of the framework are three interrelated pillars:

1. Higher-order learning outcomes: These are the core aims—what we want students to become and be able to do by the end of their education. This includes advanced discipline knowledge, cognitive skills (analysis, evaluation and synthesis), creative thinking, ethical and reflective judgement and practical problem-solving abilities. It also encompasses professional skills like teamwork, communication, and digital literacy (now including AI literacy). These outcomes reflect the "capacities to understand and act critically and creatively" that define higher education's purpose. They are often articulated in graduate profiles and programme learning goals.

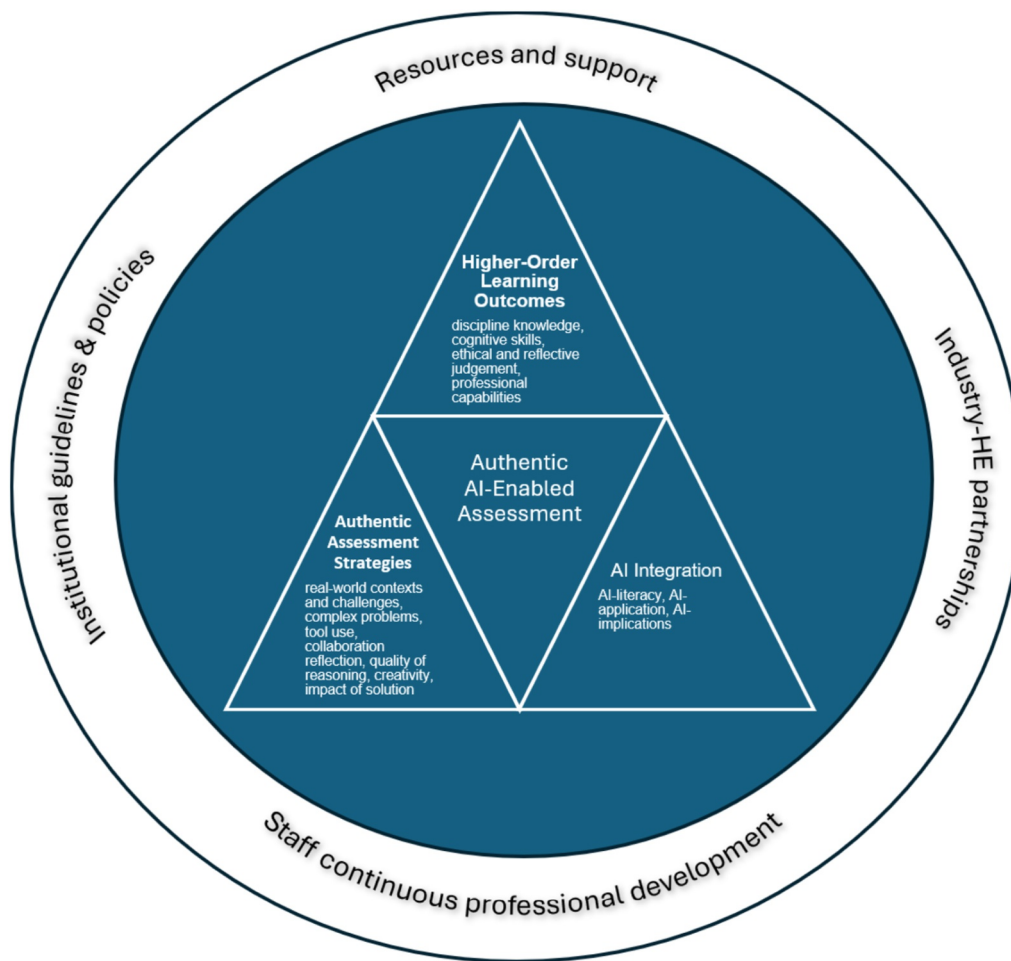


FIGURE 1 | The Authentic AI Enabled Assessment (AAA) model.

2. Authentic assessment strategies: These are the methods and task designs through which students demonstrate the above outcomes. Authentic assessments operationalise higher-order outcomes into concrete learning activities and evidence. In this framework, authentic assessments are deliberately designed to:

- mirror real-world contexts and challenges (e.g., a marketing campaign project for a real client, an investment analysis with live data, a business simulation game, etc.).
- require students to apply multidisciplinary knowledge and critical thinking to open-ended problems.
- allow the use of relevant tools and collaboration, thus focusing on process and decision-making rather than solely on the product of the output.
- include reflective components to capture ethical reasoning and learning from experience.
- be evaluated with criteria that emphasise the quality of reasoning, creativity and impact of the solution, not just the accuracy of the content.

These strategies align with assessment for learning and assessment as learning (Earl 2012), not solely assessment of learning. In other words, the assessment tasks themselves become powerful learning experiences that develop the desired competencies while also evaluating them.

3. AI utilisation and literacy: The framework integrates AI not as a threat but as a component within authentic

assessments. This pillar recognises AI as a tool in professional practice and, thus, a relevant part of the learning environment. The AAA framework positions AI as integral to learning and assessment by embedding a three-stage developmental progression: AI literacy, AI application and AI implications within authentic task design. Under this model, students collaborate with, critique and adapt AI outputs while exercising professional judgement in AI mediated environments.

Key ideas include the following:

- Teaching students when and how to use AI tools responsibly to enhance their work (e.g., using AI for brainstorming ideas, conducting preliminary research and generating data visualisations) and when not to rely on AI (recognising tasks that require human judgement or creative originality).
- Setting clear guidelines for AI use in assessments (transparency, proper citation of AI-assisted content and adherence to ethical use policies), alongside providing iterative feedback opportunities, so that the use of AI is not hidden but rather an explicit learning outcome to be assessed.
- Designing some tasks that explicitly involve AI literacy—for instance, comparing human versus AI-generated solutions, critiquing AI outputs or having students improve an AI-generated draft. This turns AI

into a pedagogical tool: Students learn by correcting AI's mistakes or identifying its biases, thereby sharpening their critical thinking.

- o Safeguarding authenticity and integrity by including a balance of controlled and noncontrolled assessments or steps within the assessment. For example, a student might use AI to help with data analysis in a finance project, but then they must orally defend their interpretation of the results and decision recommendations or engage in an in-person interactive oral discussion. This ensures the student has personally internalised the knowledge and can articulate it beyond what AI provided.

3.2 | How the AAA Framework Can Guide Assessment Design

To illustrate how the AAA framework can be used in practice, consider a conventional assessment in which students are asked to write an essay on a social issue such as housing insecurity, youth justice or the cost-of-living crisis. In its traditional form, this task may invite students to describe the issue, collect relevant information, summarise literature and present recommendations, but it remains vulnerable to generic responses and AI-generated text. This is particularly the case when tasks do not require students to demonstrate reasoning, judgement, or application in context. Using the AAA framework, an educator would begin by revisiting the higher-order learning outcomes the assessment is intended to develop. Rather than assessing only students' ability to describe the issue, the educator would specify outcomes such as critical analysis of evidence, ethical reasoning and the capacity to formulate contextually appropriate responses that account for competing stakeholder interests.

The second step is to redesign the task so that it becomes more authentic. Instead of asking for a generic report, students could be positioned as policy advisers preparing a briefing for a local council, community organisation or ministerial committee. They might be required to analyse a contemporary social issue in a defined setting, evaluate possible responses and justify a recommended course of action for a specific audience. To make the task more process-oriented, the assessment could be scaffolded into stages such as problem definition, annotated evidence scan, draft recommendation and final briefing. A reflective commentary can be included to require students to explain how their thinking evolved, the trade-offs they confronted and how they responded to feedback.

The third step is to deliberately design for AI utilisation and literacy rather than treating AI as either invisible or prohibited. Students may use AI in early stages to generate lines of inquiry, identify stakeholder groups, or produce provisional summaries of policy options. However, the task would require them to critically evaluate the AI output for factual accuracy, bias, missing perspectives, weak reasoning and the potential marginalisation of vulnerable communities. Students are assessed not only on the final quality of the briefing but also on how responsibly and critically they engaged with AI.

The fourth step is to build in controlled checkpoints that ensure students can demonstrate learning beyond what AI can produce

independently. These checkpoints enable student learning to be evidenced more directly under defined conditions. For example, after submitting the written briefing, students might participate in a short live policy hearing, viva or interactive discussion in which they must defend their recommendations, respond to questions from peers and adapt to an unexpected twist such as a sudden funding cut, a new stakeholder objection or emerging evidence that changes the policy context. This makes students' judgement and adaptability visible in real time.

The final step is to align the marking criteria with the qualities the assessment is intended to value. Rather than allocating most of the marks for the content of the written briefing, the rubric would emphasise the quality of analysis, the strength of justification, the use and critique of evidence, ethical and social awareness, the transparency and appropriateness of AI use and the student's ability to explain and defend decisions. This shifts the assessment from a task that AI can readily replicate to one that values process, reflection, professional judgement and human-AI collaboration.

3.3 | Alignment With Educational Frameworks and Principles

This conceptual model is underpinned by established educational frameworks that emphasise a tight alignment among learning outcomes, instructional methods and assessment practices. Drawing on the principle of constructive alignment (Biggs and Tang 2011) and authentic assessment (Gulikers et al. 2004), the framework ensures that the skills we value—ranging from AI literacy to ethical reasoning—are directly embedded in both how we teach and how we assess learning. Complementing this, the Assessment 2020 propositions (Boud and Associates 2010) advocate for assessments that not only measure learning outcomes but also promote self-regulation and reflective practice, thereby providing rich evidence of student growth and deeper cognitive development.

The AAA framework extends established assessment models by responding directly to the radically transformed conditions of learning created by AI. For example, the five-dimensional model reported by Gulikers et al. (2004) focuses on authenticity of task, context, criteria and performance based on professional practice, whereas Assessment 2020 propositions by Boud and Associates (2010) emphasise programme level coherence, evaluative judgement and assessment for learning. However, neither framework anticipates how AI reshapes knowledge practices, decision-making or the nature of authenticity itself. This integrated framework thus supports the development of graduates who are not only proficient in their disciplines but are also prepared to tackle real-world challenges in an AI-enhanced environment.

The conceptual framework is not a panacea. It assumes certain enabling conditions such as manageable class sizes or support for large classes, training for assessors and the availability of appropriate technology. Nonetheless, as a conceptual contribution, it provides a structured framework for transforming assessment so that it remains true to the core ideals of higher education while adapting to the rapidly evolving technological context.

4 | Discussion

Implementing the AAA framework requires coordinated attention to assessment design, delivery, policy and staff and student preparedness across programmes.

4.1 | Rethinking Assessment Design and Delivery

One immediate implication of adopting authentic AI-enabled assessment is that educators must redesign many existing assessments. This is an investment of time and effort. Institutions will need support (through professional development and possibly curriculum design assistance) to create robust authentic tasks and accompanying rubrics. Institutions might establish communities of practice where early adopters share their redesigned assessments and outcomes. Over time, a repository of exemplars can ease the burden. Importantly, these efforts are most effective when assessment redesign happens at the programme level rather than in isolation within individual courses, allowing controlled and noncontrolled assessments to be coherently sequenced across the degree.

With the AAA approach, the assessment delivery may need to adapt. For instance, more presentations, vivas, and interactive assessments will require scheduling and perhaps additional resources (like panel evaluators or technology for remote presentations). Scalability remains a challenge, particularly in large or resource-constrained contexts. Solutions include using group presentations or debates to accommodate more students in the same timeframe, leveraging teaching assistants or using asynchronous recorded group video presentations. Approaches such as online simulations and peer evaluation can further support scalability, although continued innovation is needed. Recent advances in video assessment tools and even AI-assisted grading for presentations can alleviate workload; however, careful oversight is required to ensure fairness and mitigate potential algorithmic bias.

Feedback and iterative development are an integral part of authentic assessments. Educators should plan for formative checkpoints where students get feedback on their progress (e.g., oral progress reports for feedback before the final project). This aligns with the principle of assessment for learning and helps students stay on track (since the process is observed). The University of Sydney's Assessment Framework explicitly calls for assessment practices that “facilitate reflection and judgement” and provide feedback students can act on—authentic tasks lend themselves to this, as they often unfold in stages. Following the University of Sydney's two-lane model, Auckland University of Technology is undertaking an institution-wide assessment redesign to embed a deliberate mix of controlled and noncontrolled assessments across degrees, ensuring coherent and ethical AI integration and strong assurance of learning.

4.2 | Policy and Academic Integrity Considerations

From a policy standpoint, universities will need to update their academic integrity policies to clarify the acceptable use of AI. Many have already done so, adding clauses about AI. The policy should define plagiarism not as mere use of AI but as

undisclosed or unethical use of AI that violates task guidelines. For example, if a reflection is meant to be personal, using AI to write it is deceptive. However, AI use may be appropriate where it is permitted by the task guidelines and students acknowledge how it was used. Clear guidelines for students on what is allowed in each assignment are crucial. Transparency is key: Some courses require an “AI usage statement” appended to submissions, where students state which AI tools were used and how. This normalises AI as a tool, akin to stating one's references, and holds students accountable for honesty.

Assessment boards and accreditation bodies will require education about the new forms of assessment to ensure they are understood and trusted. External examiners (in systems that use them) should be briefed on how authenticity and AI usage are managed. Accreditation bodies, such as AACSB, have already begun highlighting how schools address academic integrity in the context of AI—demonstrating a proactive approach with authentic assessment could be viewed positively in accreditation reviews, as evidence of innovation and rigorous standards. Other professional bodies, such as Certified Practising Accountants (CPA) assessment standards require institutions to have at least 50% of assessment activities be verifiable or invigilated. Notably, the CPA does not limit assessments solely to traditional in-person examinations but allows for a variety of authentic invigilated formats, such as case studies, interactive presentations, and monitored Q&A sessions, as long as student performance is directly observed in a controlled environment (CPA Australia 2023; Gedera 2023).

4.3 | Staff and Student Preparation

Adopting authentic assessment widely represents a culture shift for both educators and students. Educators may initially fear losing control or worry about grade inflation (e.g., if many students do well on a real-world task). Professional development and sharing evidence of efficacy can mitigate these concerns. It's important to highlight that authentic does not mean “easy”—these authentic assessment tasks can be more challenging and reveal student ability than traditional assessments. Educators may observe a wider distribution of performance, as students with well-developed skills can demonstrate their capabilities more fully, whereas others are required to engage more actively rather than relying on memorisation alone. Rigorous rubrics and possibly double marking or moderation of a sample of works can ensure grading fairness.

Students, for their part, need to be guided to succeed in authentic tasks. Many incoming students might initially be anxious when told their assessment is, say, a complex project with multiple deliverables. Scaffolding is essential: educators should break tasks into milestones, provide exemplars and explicitly teach skills like writing, presentation, collaboration or using AI tools. This is part of “learning by doing.” For example, if an assessment requires using a data visualisation tool, some class time should be devoted to training on that tool. If an ethical analysis is expected, the course should include ethical frameworks and case discussions beforehand. Essentially, the teaching must align with the assessment demands—which is just good practice, but worth emphasising.

Interestingly, developing genuine AI literacy requires far more than a single workshop for staff or students. Both educators and students need sustained hands-on opportunities to practise using AI in low-stakes, supportive environments. Creating safe spaces for experimentation and sharing of examples and where students can learn to use AI as a thinking partner rather than a shortcut helps normalise responsible and reflective AI use. Short modules introducing AI fundamentals, such as how AI works, its limitations and how to critically evaluate its outputs, remain valuable, but they must be complemented by ongoing opportunities to integrate AI into teaching and learning activities. Embedding this practice across courses ensures that, by the time students engage in major authentic assessments, both educators and students possess the baseline AI literacy, ethical understanding and confidence needed to use AI responsibly and effectively.

4.4 | Ensuring Quality and Fairness

One possible critique of authentic assessment might be about standardisation: How do we ensure fairness when students are doing different projects or using different approaches? The answer lies in robust rubrics and clear criteria. The rubric should focus on the qualities of thinking and performance rather than specific content. For instance, in a finance case analysis, criteria might include “accuracy and depth of analysis of financial data,” “logical justification of recommendations” and “consideration of ethical and risk factors” rather than expecting a single correct numeric answer. Moderation processes (having multiple graders or a second reviewer) can increase reliability. Technology can help here too: Many learning management systems (LMS) support detailed rubric grading, which can be analysed to ensure consistency across sections or graders.

A related concern is the workload for students. Programme-level coordination (programmatic approach to assessments) is important so that, for example, not every course has a major group project due in the same week. Staggering due dates and perhaps integrating tasks across courses (interdisciplinary capstones) can help. Some institutions have moved towards fewer but more integrative assessments—which students often appreciate, as it feels less like busy work and more like doing something substantial.

It’s also crucial to consider equity. Some students might have prior experience or access to resources (e.g., familiarity with certain software or personal connections to industry) that give them an advantage in authentic tasks. To mitigate this, ensure all students have access to necessary tools (provide software licenses, data sets, etc.), and avoid assessments that rely too much on resources outside a student’s reach. If a task involves working with an external client, educators might source the clients to ensure every team has a comparable opportunity rather than expecting students to find their own (unless that networking aspect is an intended learning outcome, in which case support can still be given). Additionally, if AI tools are used, ensure every student can access them—which in most cases is true for free or with institutional subscriptions and support.

4.5 | Examples of Authentic Assessments in Business Education

To illustrate concretely how authentic assessment can be implemented, below are some examples from business education. These examples are suitable across undergraduate and postgraduate levels (with complexity scaled appropriately) and each is designed to assess specific skills, knowledge, and attributes:

- **Interactive Oral Assessment**

Task: Interactive Oral Assessments are dynamic, scenario-based conversations. In this assessment, students engage in unscripted discussions—either one-on-one or in small groups—with an examiner or panel (Logan-Fleming and Sotiriadou 2020). For example, a student may be presented with a complex business scenario—such as an ethical dilemma in corporate finance or a strategic audit decision—and required to articulate their approach, justify their decisions and reflect on alternative strategies in real time. A distinctive feature of this assessment is the inclusion of an unexpected twist during the conversation. At a pre-determined point, the examiner may introduce new information, an alternative perspective or a sudden change in the scenario, compelling the student to think on their feet and quickly adjust their strategy.

Assesses: This method evaluates not only the final verbal response but also the learning process itself. It measures competencies such as oral communication, critical thinking, adaptability, and the ability to integrate interdisciplinary knowledge under pressure. The twist element is designed to assess the student’s capacity to demonstrate resilience, creative problem-solving and ethical judgement in the face of unforeseen challenges. By requiring students to engage in reflective dialogue and adapt their responses to evolving scenarios, interactive oral assessments provide a comprehensive picture of each student’s readiness for real-world professional interactions.

AI Integration: Although AI tools can support the preparatory stage—for example, by generating practice prompts, offering spoken feedback, or transcribing rehearsal sessions (AI literacy and application)—the assessment itself is intentionally designed as a live, unscripted conversation featuring an unexpected twist. In this moment, students must think independently, adapt rapidly, justify their decisions, and respond authentically without AI support, demonstrating awareness of when human judgement, ethical reasoning and situational awareness cannot be automated or delegated to AI systems (implications).

- **Consulting Project for a Real Client**

Task: Students (individually or in teams) act as consultants for a real or simulated client organisation to solve a business problem (e.g., improving a company’s social media strategy or expanding into a new market). They interact with the client (role-played if necessary), gather and analyse data and present recommendations in a professional report and interactive presentation. The project is broken into tasks (scaffolded) where students demonstrate what they have learned and receive feedback at every step of the process. Students are not only evaluated for the product but also for the process of engaging with the client and collaborating with peers.

Assesses: Problem-solving and analytical skills (through researching the problem and data analysis), application of business knowledge (integrating marketing, finance and management concepts as needed), communication skills (professional report writing and oral presentation), ethical reasoning (considering the client's stakeholders and any ethical implications of recommendations) and project management/teamwork (if done in groups).

AI Integration: Students may use AI tools to support early research, generate initial ideas or structure draft reports, but they must verify, refine and justify all AI-generated material. Through the AAA lens, students develop AI literacy by critically evaluating the accuracy, credibility and ethical risks of AI-produced insights. They then move to AI application as they strategically employ AI to accelerate data analysis, generate alternative solutions or explore stakeholder perspectives, documenting their use transparently. The assessment also prompts reflection on the implications of AI, requiring students to evaluate how AI may reproduce cultural biases, overlook marginalised or missing voices and simplify complex organisational realities. In doing so, students must apply their own ethical reasoning and human judgement to determine which insights can be trusted, which require challenge or correction, and how automated contributions should be integrated responsibly into consulting practice.

- **Business Simulation Game**

Task: Students participate in a computer-based business simulation (e.g., a simulated competitive market where teams run virtual companies) (Neely and Tucker 2012). Over several rounds, they make strategic decisions (pricing, investment, marketing spend, etc.) and adapt to market conditions and competitor moves. They are assessed on a combination of their company's performance in the simulation and an interactive debriefing at the end.

Assesses: Strategic thinking and decision-making (applying theory to dynamic decisions), data analysis (interpreting financial and market reports generated each round), adaptability (responding to unexpected outcomes in the simulation), teamwork (if teams are used) and reflective learning (interactive debriefing on what they learned from successes/failures and how theory informed their strategy).

AI Integration: Although AI is not embedded in most simulation platforms, students may experiment with AI tools to forecast possible outcomes, organise information or test strategic options. This creates a natural space for students to build AI literacy as they observe where predictive AI succeeds and where it falls short in dynamic, uncertain environments. Students then engage in AI application, using AI to support aspects of analysis while comparing AI-generated suggestions with their own reasoning. Reflection on the implications of AI is central because students evaluate when AI can enhance strategic decision-making, when reliance on AI may be misleading and how human insight remains critical when competing against unpredictable human teams.

- **Entrepreneurial Venture Pitch**

Task: Students develop a concept for a new venture (product or service), create a business plan or business model

canvas and pitch it to a panel of judges (which could include educators, peers, or actual entrepreneurs). Deliverables might include a slide deck, a written executive summary and an oral pitch.

Assesses: Creativity and innovation (coming up with a viable and novel business idea), market research and analysis (justifying the opportunity with real market data), financial literacy (basic financial projections in the plan), communication and persuasion (the effectiveness of the pitch) and resilience/feedback response (often judges Q&A is included, testing how students think on their feet and respond to critique).

AI Integration: Students may use AI to brainstorm venture ideas, draft parts of the business plan or simulate pitch practice, but all market data, financial assumptions and strategic claims must be independently validated. Applying the AAA framework, students cultivate AI literacy by recognising the capabilities and limitations of AI in entrepreneurship, particularly regarding originality and data reliability. They then demonstrate AI application by using AI to prototype materials, refine messaging or explore alternative business models. The assessment also highlights the implications of AI, prompting students to consider ethical concerns (e.g., authenticity, plagiarism, bias), the shifting nature of entrepreneurial work and the importance of maintaining human creativity and judgement when developing new ventures.

- **Ethical Decision-Making Role-Play**

Task: Students are given a case study of an ethical dilemma in business (e.g., a conflict-of-interest scenario, a whistleblowing situation, a supply chain labour issue). In a live session, each student is assigned a role (e.g., CEO, HR manager, affected employee, etc.) and must discuss or negotiate a resolution from their perspective. Alternatively, it can be a simulation of an internal meeting or a stakeholder consultation. After the role-play, students individually submit a reflection or an ethics audit report on what the company should do.

Assesses: Ethical reasoning and values (through the positions they argue and the final recommendations they make, judged against ethical frameworks taught in class), oral communication and negotiation (during the role-play interaction), perspective-taking (ability to argue from an assigned role even if it conflicts with personal view, indicating an understanding of multiple stakeholders) and critical thinking (balancing business interests with ethical principles).

AI Integration: AI may support the preparatory phase by helping students explore ethical frameworks, summarise relevant regulations or outline stakeholder arguments, but the live role-play requires students to internalise these ideas and rely on their own judgement. Through the AAA framework, students develop AI literacy as they assess the adequacy and ethical risks of AI-generated arguments. They then engage in AI application by using AI to clarify concepts or test reasoning before the activity. In the role-play and subsequent reflection, students examine the implications of AI, recognising that ethical decision-making cannot be outsourced to algorithms and that professional responsibility requires human reasoning, empathy and values-based judgement.

- Integrated Reflective e-Portfolio

Task: Throughout the programme, students accumulate evidence of their skill development and learning achievements in an e-portfolio. This may include work samples from various courses (presentations, simulations, projects etc.), reflections on internships or work placements, feedback from peers or mentors and a final reflective video connecting their academic learning to their personal and professional growth. At the end of their degree (or at the end of the year), they submit the portfolio for assessment, possibly coupled with an exit interview or presentation before moving to the next level.

Assesses: Reflective thinking and self-assessment (through the quality of reflections on experiences—for example, how did they handle a team conflict in a project and what did they learn), integrative learning (ability to connect insights across courses and experiences, showing a holistic development), communication (the organisation and presentation of the portfolio for a professional audience) and lifelong learning orientation (evidenced by how they plan future learning goals in their reflection).

AI Integration: Students may use AI tools to enhance the presentation of their portfolio or prompt deeper reflection, such as through critique or question-generation, but the narrative must remain authentically their own. Through AAA, students first develop AI literacy by understanding how AI can influence reflective writing and by identifying inappropriate or overly mechanistic use. They then put AI application into practice by using AI to strengthen clarity, structure or critical depth while maintaining ownership of the ideas. The portfolio also encourages reflection on the implications of AI, as students assess how AI has shaped their learning journey, evaluate the ethical boundaries of AI-supported reflection and consider the role of AI in ongoing professional development.

Each of these examples demands high-level skills, yields rich and individualised evidence of learning and, in most cases, can constructively incorporate AI use. Although the examples above are from business education, the AAA framework is readily applicable across other fields. For example, in history or cultural studies, students might use AI to generate a draft historical narrative based on primary sources (AI literacy and application), then evaluate the narrative for accuracy, perspective, missing voices and ethical implications such as bias, ethical, social and cultural aspects (AI implications) before presenting it to the class. This requires students to interrogate AI outputs using disciplinary expertise and ethical reasoning, ensuring that AI enhances rather than replaces human intellectual labour while also making their own analytical and communicative skills clearly visible.

5 | Conclusion and Future Directions

The disruption wrought by AI in higher education has provoked understandable anxiety while also reinforcing the need to reconsider whether existing assessment practices meaningfully reflect the educational mission of higher education. By offering the AAA framework, this article provides a way for higher education to reconceptualise assessment for an AI-enabled

world—emphasising human–AI collaboration, the staged development of AI literacy and the design of authentic tasks that strengthen students' capacity for judgement, contextualised reasoning and responsible engagement with AI. Rather than positioning AI solely as a threat, it can be understood as a catalyst for enhancing the relevance and integrity of assessment in contemporary contexts. In this way, authentic AI-enabled assessment supports closer alignment between teaching, learning and the capabilities that higher education seeks to develop.

In an AI-enabled context, authentic assessment also offers an adaptable and forward-looking approach. By integrating AI literacy into our learning outcomes and assessment methods, we ensure our graduates are not left behind in a world where AI is ubiquitous. Instead of banning AI, we teach students how to leverage AI responsibly—when to trust it, when to doubt it and how to add the human judgement that AI lacks. It also reflects the trust that universities place in students as emerging professionals: We trust them enough to give them real challenges and real tools, guiding them to become trustworthy in the process. Realising this shift will require sustained effort from students themselves, educators, institutions and accrediting bodies.

There remains a need for empirical research to evaluate the impact of these approaches on student learning, outcomes, and experiences. Although early evidence suggests improvements in engagement, further longitudinal and programme-level studies are needed to examine how controlled checkpoints, AI literacy development and authentic task design interact over time to influence student capability.

More broadly, the integration of authentic assessment and AI may influence the wider educational ecosystem, including admissions and credentialing practices. Portfolios and micro-credentials, for example, may play an increasing role in recognising authentic demonstrations of capability and supporting lifelong learning beyond formal qualifications.

Higher education can emerge stronger, with graduates who are not only knowledgeable but also agile, ethically grounded and adept at using advanced tools to solve real problems. Authentic assessment is the vehicle for this transformation—connecting the enduring purpose of higher education with the demands of the present and future.

Author Contributions

Dilani Gedera: conceptualization, writing – original draft, writing – review and editing, methodology, validation, visualization, investigation.

Acknowledgments

The author has nothing to report.

Funding

The author has nothing to report.

Conflicts of Interest

The author declares no conflicts of interest.

Data Availability Statement

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

- Ashford-Rowe, K., J. Herrington, and C. Brown. 2014. "Establishing the Critical Elements That Determine Authentic Assessment." *Assessment & Evaluation in Higher Education* 39, no. 2: 205–222. <https://doi.org/10.1080/02602938.2013.819566>.
- Barnett, R. 1997. *Higher Education: A Critical Business*. Open University Press.
- Biggs, J., and C. Tang. 2011. *Teaching for Quality Learning at University*. 4th ed. Open University Press.
- Black, P., and D. Wiliam. 2009. "Developing the Theory of Formative Assessment." *Educational Assessment, Evaluation and Accountability* 21, no. 1: 5–31. <https://doi.org/10.1007/s11092-008-9068-5>.
- Boud, D., and Associates. 2010. *Assessment 2020: Seven Propositions for Assessment Reform in Higher Education*. Australian Learning and Teaching Council. https://www.kpu.ca/sites/default/files/Teaching%20and%20Learning/Assessment_2020_final%5B1%5D.pdf.
- Cotton, D. R. E., P. A. Cotton, and J. R. Shipway. 2024. "Chatting and Cheating: Ensuring Academic Integrity in the Era of ChatGPT." *Innovations in Education & Teaching International* 61, no. 2: 228–239. <https://doi.org/10.1080/14703297.2023.2190148>.
- CPA Australia. 2023. By-Laws. <https://www.cpaaustralia.com.au/-/media/project/cpa/corporate/documents/about-cpa/governance/by-laws/unnamed-item.pdf>.
- Dwivedi, Y. K., L. Hughes, E. Ismagilova, et al. 2021. "Artificial Intelligence (AI): Multidisciplinary Perspectives on Emerging Challenges, Opportunities, and Agenda for Research, Practice and Policy." *International Journal of Information Management* 57: 101994. <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>.
- Earl, L. M. 2012. *Assessment as Learning: Using Classroom Assessment to Maximise Student Learning*. 2nd ed. Corwin Press.
- Evangelista, E. D. L. 2025. "Ensuring Academic Integrity in the Age of ChatGPT: Rethinking Exam Design, Assessment Strategies, and Ethical AI Policies in Higher Education." *Contemporary Educational Technology* 17, no. 1: ep559. <https://doi.org/10.30935/cedtech/15775>.
- French, S., A. Dickerson, and R. A. Mulder. 2024. "A Review of the Benefits and Drawbacks of High-Stakes Final Examinations in Higher Education." *Higher Education* 88, no. 3: 893–918. <https://doi.org/10.1007/s10734-023-01148-z>.
- Gedera, D. 2023. "A Holistic Approach to Authentic Assessment." *Asian Journal of Assessment in Teaching and Learning* 13, no. 2: 23–34. <https://doi.org/10.37134/ajatel.vol13.2.3.2023>.
- Gonzales, N. A. P. 2020. "21st Century Skills in Higher Education: Teaching and Learning at Ifugao State University, Philippines." *Asian Journal of Assessment in Teaching and Learning* 10, no. 2: 72–81. <https://doi.org/10.37134/ajatel.vol10.2.8.2020>.
- Gulikers, J. T. M., T. J. Bastiaens, and P. A. Kirschner. 2004. "A Five-Dimensional Framework for Authentic Assessment." *Educational Technology Research and Development* 52, no. 3: 67–86. <https://doi.org/10.1007/BF02504676>.
- Holmes, W., and I. Tuomi. 2022. "State of the Art and Practice in AI in Education." *European Journal of Education* 57, no. 4: 542–570. <https://doi.org/10.1111/ejed.12533>.
- Ishmuradova, I. I., S. P. Zhdanov, S. V. Kondrashev, N. S. Erokhova, E. E. Grishnova, and N. Y. Volosova. 2025. "Pre-Service Science Teachers' Perception on Using Generative Artificial Intelligence in Science Education." *Contemporary Educational Technology* 17, no. 3: ep579. <https://doi.org/10.30935/cedtech/16207>.
- Lodge, J., S. Howard, M. L. Bearman, et al. 2023. *Assessment Reform for the Age of Artificial Intelligence*. Tertiary Education Quality and Standards Agency. <https://www.teqsa.gov.au/sites/default/files/2023-09/assessment-reform-age-artificial-intelligence-discussion-paper.pdf>.
- Lodge, J. M. 2024. *The Evolving Risk to Academic Integrity Posed by Generative Artificial Intelligence: Options for Immediate Action*. Tertiary Education Quality and Standards Agency. <https://www.teqsa.gov.au/site/s/default/files/2024-08/evolving-risk-to-academic-integrity-posed-by-generative-artificial-intelligence.pdf>.
- Logan-Fleming, D., and P. Sotiriadou. 2020. *Interactive Oral Assessments: A Viable Solution to Mitigate Potential Reverse Effects of Global Threats to the Higher Education Sector*. Griffith University. <https://sway.cloud.microsoft/Q2s0Bm3ILkWTGII?ref>.
- Mueller, J. 2005. "The Authentic Assessment Toolbox: Enhancing Student Learning Through Online Faculty Development." *Journal of Online Learning and Teaching* 1, no. 1: 1–7. https://jolt.merlot.org/vol1_no1_mueller.htm.
- Neely, P., and J. Tucker. 2012. "Using Business Simulations as Authentic Assessment Tools." *American Journal of Business Education* 5, no. 4: 449–456. <https://doi.org/10.19030/ajbe.v5i4.7122>.
- Nussbaum, M. C. 2010. *Not for Profit: Why Democracy Needs the Humanities*. Princeton University Press.
- Perrotta, C., and N. Selwyn. 2020. "Deep Learning Goes to School: Toward a Relational Understanding of AI in Education." *Learning, Media and Technology* 45, no. 3: 251–269. <https://doi.org/10.1080/17439884.2020.1686017>.
- Rudolph, J., S. Tan, and S. Tan. 2023. "ChatGPT: Bullshit Spewer or the End of Traditional Assessments in Higher Education?" *Journal of Applied Learning and Teaching* 6, no. 1: 342–363. <https://doi.org/10.37074/jalt.2023.6.1.9>.
- Singh, P., R. X. Thambusamy, and M. A. Ramly. 2014. "Fit or Unfit? Perspectives of Employers and University Instructors of Graduates' Generic Skills." *Procedia-Social and Behavioral Sciences* 123: 315–324. <https://doi.org/10.1016/j.sbspro.2014.01.1429>.
- Villarroel, V., D. Boud, S. Bloxham, D. Bruna, and C. Herrera-Seda. 2018. "Authentic Assessment: Creating a Blueprint for Course Design." *Assessment & Evaluation in Higher Education* 43, no. 5: 840–854. <https://doi.org/10.1080/02602938.2017.1412396>.
- Vlachopoulos, D., and A. Makri. 2024. "A Systematic Literature Review on Authentic Assessment in Higher Education: Best Practices for the Development of 21st Century Skills, and Policy Considerations." *Studies In Educational Evaluation* 83: 101425. <https://doi.org/10.1016/j.stueduc.2024.101425>.
- Wiggins, G. 1990. "The Case for Authentic Assessment." *Practical Assessment, Research and Evaluation* 2, no. 1: 2. <https://doi.org/10.7275/fbf1-mm19>.
- World Economic Forum. 2020. *The Future of Jobs Report 2020*. https://www3.weforum.org/docs/WEF_Future_of_Jobs_2020.pdf.