Studies in Higher Education

ISSN: 0307-5079 (Print) 1470-174X (Online) Journal homepage: www.tandfonline.com/journals/cshe20

Walking the tightrope of quality assessment: balancing perspectives and priorities of stakeholder groups

Sandris Zeivots, Sue Wright, Lynne Harris, Andrew Cram, Elaine Huber, Corina Raduescu, Amanda White & Andrew Brodzeli

To cite this article: Sandris Zeivots, Sue Wright, Lynne Harris, Andrew Cram, Elaine Huber, Corina Raduescu, Amanda White & Andrew Brodzeli (26 Jan 2026): Walking the tightrope of quality assessment: balancing perspectives and priorities of stakeholder groups, Studies in Higher Education, DOI: 10.1080/03075079.2026.2619909

To link to this article: https://doi.org/10.1080/03075079.2026.2619909

Published online: 26 Jan 2026.

STUDIES IN HIGHER EDUCATION
https://doi.org/10.1080/03075079.2026.2619909

Routledge
Taylor & Francis Group

Walking the tightrope of quality assessment: balancing perspectives and priorities of stakeholder groups

Sandris Zeivots [a], Sue Wright [b], Lynne Harris [c], Andrew Cram [a], Elaine Huber [a], Corina Raduescu [a], Amanda White [b] and Andrew Brodzeli[d]

[a]Business School, The University of Sydney, Sydney, Australia; [b]Business School, University of Technology Sydney, Sydney, Australia; [c]Sydney Metropolitan Institute of Technology, Sydney, Australia; [d]The University of Sydney, Sydney, Australia

ABSTRACT

The ongoing evolution of digital technologies, particularly Generative Artificial Intelligence, continues to shape and challenge assessment design in higher education. Given the complex and sometimes competing factors that contribute to assessment design, and the evolving digital landscape in which assessment is placed, this study examines the perspectives and priorities of five key stakeholder groups – educators, students, employers of graduates, accrediting bodies, and institutional policy-makers – regarding the defining characteristics of quality assessment. Using a mixed-methods approach, we conducted interviews, focus groups, and a national survey to extend a framework for designing quality digital assessments in business education that was originally developed using educator perspectives only. The findings highlight the importance of balancing academic integrity, feedback quality, student experience, and authenticity in assessment design to address stakeholder perspectives. They also extend the framework by including two additional design elements: purpose and technology, and by emphasising the value of dialogue about contrasting interpretations of assessment quality. The study provides a refined framework that incorporates nuanced differences in stakeholder priorities, supports educators in designing digital assessments that respond to stakeholder needs, and encourages co-design and shared accountability.

ARTICLE HISTORY
Received 14 May 2025
Accepted 16 January 2026

KEYWORDS
Digital assessment; business education; assessment design; Generative Artificial Intelligence; higher education

Introduction

Higher education has been transformed by advances in digital technologies in the last two decades. As Nieminen, Bearman, and Ajjawi (2023) note, there is little assessment in higher education that is not at least partially digital, a trend that accelerated during the COVID-19 pandemic and with the arrival of Generative Artificial Intelligence (genAI). Technologies such as genAI can enhance assessment design through automated feedback, personalised learning, and data analytics (Lodge et al. 2023). GenAI also presents new challenges, particularly in maintaining academic integrity, providing equitable access, ensuring ethical use and requiring effective training. In this context, the ongoing refinement of digital assessment strategies is essential to prepare students for the changing workplace and meet evolving stakeholder demands for relevant and ethical assessment.

CONTACT Sandris Zeivots sandris.zeivots@sydney.edu.au Business School, The University of Sydney, H70 Belinda Hutchinson Building, Sydney, NSW 2006, Australia

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

This study builds on the growing body of research into digital assessment (see Bearman, Nieminen, and Ajjawi 2023; Grosseck et al. 2024; Swiecki et al. 2022), specifically extending prior research that developed and evaluated a framework for quality online assessments in business education (Huber et al. 2024). The original context of business education, while vocationally oriented, lacks mandatory work-integrated learning placements and therefore has characteristics similar to generic qualifications in arts and sciences. As such, the framework offers potential relevance across a broader range of higher education contexts. With input from educators, Huber et al. (2024) identified a need for authentic assessment practices that enhance student engagement while ensuring academic integrity and scalability. Challenges faced by educators included increased time and effort required for assessment design, technology access, and changes in feedback mechanisms (Cram et al. 2022).

In the current study, we evaluate the validity of the framework introduced by Huber et al. (2024) using a new dataset that incorporates perspectives from five stakeholder groups: educators, students, employers of graduates, accrediting bodies and institutional policy-makers. While educators and students are directly involved in assessment, other groups provide critical and complementary perspectives in the oversight of assessment in business education or the provision of post-tertiary experience.

Digital assessments

Digital assessment refers to evaluating student performance, capabilities and learning using digital tools and platforms (Grosseck et al. 2024). These assessments offer scalability, flexibility, automated feedback, student engagement, and personalised learning experiences (Heil and Ifenthaler 2023; Timmis et al. 2016; Viberg et al. 2024). However, digital assessments also pose challenges related to academic integrity, authenticity and skill assessment (Fawns et al. 2025; Nieminen, Bearman, and Ajjawi 2023).

Digital assessments frequently mimic traditional forms like exams, quizzes, reports and presentations (Cram et al. 2022; Heil and Ifenthaler 2023). Bearman, Nieminen, and Ajjawi (2023) argue that further transformation is needed in how students' performances and capabilities are elicited, evaluated, and improved. To fulfil their potential, digital assessments should enhance students' experience, develop digital literacies and foster human capabilities required for a digitally-integrated world.

Several frameworks guide the design of digital assessments in higher education. The Framework for Authentic Assessment (Gulikers, Bastiaens, and Kirschner 2004) focuses on creating assessments that mirror real-world tasks and challenges to engage students in applying knowledge to practical contexts and developing critical thinking, problem solving and evaluative skills. The Assessment Design Decisions Framework (Bearman et al. 2016) aims to support academics to design assessments customised to their courses and highlights decisions relating to six areas: purpose, context, learner outcomes, tasks, feedback processes and interactions. Both frameworks recognise the importance of different stakeholders and contexts in shaping assessment-related experiences.

Digital assessments offer scalability to accommodate large cohorts with flexible, asynchronous tasks (Viberg et al. 2024), while genAI-powered tools offer timely feedback to help students identify areas for improvement (Mate and Weidenhofer 2022). Furthermore, they offer opportunities for students to represent knowledge, either individually or collectively, and engage in complex decision-making and problem-solving (Timmis et al. 2016).

Despite these benefits, digital assessments raise concerns about academic integrity (Bearman, Nieminen, and Ajjawi 2023). The risk of impersonation, outsourcing and inappropriate assistance, including the use of genAI tools like ChatGPT to produce assessment responses, challenges the validity of assessment tasks (Cotton, Cotton, and Shipway 2024). Institutions are encouraged to implement guidelines to maintain academic standards (Swiecki et al. 2022), although this remains an ongoing challenge.

The design of digital assessments requires diverse stakeholder input (Gonsalves and Lin 2025). For example, student insights are essential for evaluating the effectiveness of digital tools, task clarity,

and feedback relevance, helping to design assessments that are fair, engaging and support student learning (Bennett et al. 2017). Educators are pivotal in ensuring digital assessments align with curriculum objectives and provide feedback on their practicality and feasibility (Bearman, Nieminen, and Ajjawi 2023). Additionally, institutional policy-makers align assessments with institutional and regulatory requirements (TEQSA 2024).

This study extends the existing digital assessment frameworks by embedding stakeholder-specific insights relevant to business education. Engaging diverse stakeholders in the development of assessment design frameworks enhances their relevance, legitimacy, and uptake across higher education. Similar principles apply in co-design processes (Zeivots et al. 2025), where multiple stakeholders collectively contribute to curriculum design, including assessment tasks. Involving academic staff, professional educators, students, employers and quality assurance personnel in design processes can surface broader needs, support shared ownership, and strengthen a framework's applicability (Hopfenbeck et al. 2023). Such inclusive approaches respond to the increasing complexity of digital and authentic assessment, ensuring that frameworks reflect current practice and future challenges (Martel and Pérez Garcias 2024; Nieminen, Bearman, and Ajjawi 2023).

Digital assessment: design considerations and contextual factors

Huber et al. (2024) examined business educators' perspectives of key elements in the design and evaluation of quality digital assessments and developed a framework consisting of six design considerations alongside four contextual factors (Table 1). The six design considerations identify elements that are considered and appropriately weighted during assessment design, because actions to improve quality in one area frequently involve a trade-off against reduced quality in another. Ajjawi et al. (2024) call this 'satisficing', a process of making optimal design decisions under different, sometimes incompatible requirements. An assessment task meeting all criteria is a unicorn, although educators may achieve balance across assessments within a course. The design considerations apply whether the assessments have primarily formative or summative

Table 1. Design considerations and contextual factors for evaluating and designing quality digital assessments.

Design consideration	Description	Related literature
Academic integrity	Security of the assessment, including assuring against impersonation, outsourcing and inappropriate assistance.	Birks et al. 2020; Hancock et al. 2023.
Student experience	Maximises convenience, comfort and ease of concentration; minimises technical disruptions, anxiety, stress and excessive cognitive load.	Cramp et al. 2019; Butler-Henderson and Crawford 2020.
Authenticity	Reflection of skills and knowledge common to professional settings.	Sotiriadou et al. 2020; Fawns et al. 2025; Nieminen, Bearman, and Ajjawi 2023.
Information integrity	Protection of student information and data, including personal data and assessment content.	Kharbat and Abu Daabes 2021; Okada et al. 2019.
Quality feedback	Timely, encourages educational dialogue and allows students to understand and improve their performance.	Dawson and Henderson 2017; Winstone and Boud 2022.
Equity of access	Equal opportunities for student success by removing technical and logistical barriers, avoiding potential discrimination, and enabling customisations for individual student needs.	Stephenson and Harvey 2022; Tai et al. 2023.
Contextual factors		
Scale of delivery	Different techniques for collection of assessment data, grading, feedback, and management/administration suit different student cohort sizes.	Cavalcanti et al. 2021; Dawson and Henderson 2017.
Resources	Available staff time and budget to develop, implement and mark assessments, as well as enabling technologies and space.	Cramp et al. 2019.
Institutional policies	Influence how assessments should (or should not) be designed.	Birks et al. 2020; Chan 2023.
Accreditation requirements	Additional conditions for assessment design imposed by professional associations or regulatory bodies.	Ali, Narayan, and Gedera 2022; Hancock et al., 2023.

goals, although the relative importance of each element may differ accordingly. The four contextual factors modulate the parameters of assessment design and evaluation in business education and highlight why the enactment of quality assessment varies between contexts.

This framework examines assessment quality criteria at a more conceptual level relative to other assessment design frameworks anchored in technical characteristics of assessment tasks (e.g. Gulikers, Bastiaens, and Kirschner 2004) or the design process (e.g. Bearman et al. 2016). It is intended to support assessment (re)design, evaluation, validation and dialogue, including helping justify assessment design to an accrediting body (Huber et al. 2024).

Changing context: GenAI

The accelerating evolution of genAI has intensified long-standing pressures and limitations on assessment design in higher education (Bhullar, Joshi, and Chugh 2024) and challenges foundational assumptions about what constitutes valid evidence of learning and students' own work. Developments in genAI have also transformed workplaces and everyday practices. Educators are therefore expected to incorporate digital technology, including genAI, while also ensuring that assessment provides credible evidence of student learning. Supporting these expectations, there have been guidelines and calls that emphasise adapting assessment design to genAI capabilities, revising policy and building AI literacies. In the United Kingdom, Universities UK and Jisc have released sector guidance to adapt assessment design to consider genAI capabilities. In the United States, the Council of Chief Academic Officers and EDUCAUSE have called for systemic responses to genAI. In Australia, the higher education regulator, Tertiary Education Quality and Standards Agency (TEQSA) commissioned papers outlining principles and strategies for assessment reform in response to AI (Lodge et al. 2023) and established a practice hub to support the use of genAI in higher education curricula (TEQSA 2024). These responses recognise that genAI is more than 'simply another tool' to be accommodated, it reconfigures what counts as knowledge, how it is produced and how it is used for curriculum design purposes (Rapanta et al. 2025; Zeivots et al. 2025).

Moorhouse et al.'s (2023) review of assessment guidelines from 50 global higher education institutions highlights the need to redesign tasks that draw on creativity, critical thinking, contextual knowledge and authentic, practice-based scenarios that are less amenable to production by genAI (Salinas-Navarro et al. 2024). They emphasise assessing the process and stages of assessment preparation and recommend deliberately incorporating genAI within tasks where appropriate. In addition, they note the value of assessment types in which individual identity can be verified and authenticated, such as debates, oral presentations, and invigilated exams. These proposals align with wider calls to develop pedagogically driven frameworks that articulate acceptable AI use, connect it to learning outcomes and support assessment redesign (Corbin et al. 2025; Perkins, Roe, and Furze 2025).

At the same time, research shows that genAI unsettles assumptions about assessment literacy, authorship and feedback for both students and educators. Many students already use genAI to generate ideas, interpret tasks and refine drafts, yet their ability to judge the reliability, limitations and epistemic status of AI-generated material is uneven (Henderson et al. 2025; Walton et al. 2025). This reflects uncertainty about acceptable assistance, and how much critical engagement is expected of the student. Educators also face new interpretive demands in distinguishing student thinking from AI influence, and in supporting responsible use that enhances rather than obscures learning (Crawford, Cowling, and Allen 2023; Farazouli et al. 2025).

Despite these challenges, little is known about how stakeholder priorities converge, or conflict, in shaping assessment frameworks. It is timely to consider how diverse stakeholder priorities inform the (co-)design of quality assessment and to assess the ongoing utility of Huber et al. (2024) framework for assessment design. The study addresses this gap in two phases: first, by analysing the perspectives of five key stakeholder groups concerning the factors that influence the quality of digital

assessment in business education (Phase 1), and second, by examining how these priorities align with, and potentially modify, the framework (Phase 2).

Methodology

The study was approved by the University of Sydney Ethics Committee (Approval No.: 2021/800). The perspectives and priorities of the five targeted stakeholder groups (educators, students, employers of graduates, accrediting bodies and institutional policy-makers) were examined using interviews or focus groups (Phase 1) and an online survey (Phase 2), where the analysis from Phase 1 informed the development of Phase 2.

Phase 1. Focus groups and interviews

Participants and recruitment

A total of 46 participants from the five stakeholder groups were recruited via: (1) announcements on LinkedIn and through professional societies such as HERDSA and ASCILITE (educators; $n = 12$); (2) emails to heads of Australian learning and teaching units, Pro Vice-Chancellors, Associate Deans and Program Coordinators (institutional policy-makers; $n = 5$); (3) emails to accrediting bodies (accrediting body representatives; $n = 6$); (4) emails to addresses provided by institutional community engagement offices (employers of graduates; $n = 8$); and (5) notices distributed through student societies and online learning information hubs (students; $n = 15$). Participants engaged in online interviews ($n = 10$) or one of seven online focus groups ($n = 36$).

Materials

Participants answered a series of open-ended questions about each design consideration and contextual factor, focusing on its importance in the current assessment context. Participants were also asked if anything additional had been missed.

Questions specific to stakeholder groups were included. For example, employers of graduates and accreditation body representatives were asked to reflect on the relationship between quality digital assessment and the workplace. Institutional policy-makers were asked about modifications to institutional policies in response to genAI and its impact on high-quality digital assessment. Students were asked about the usefulness of different assessment types.

Analysis

The qualitative data was analysed using thematic analysis (Nowell et al. 2017), specifically in relation to the framework parameters and contextual factors, and areas for further consideration. We ensured the rigour and trustworthiness of our analysis using Nowell et al.'s (2017) criteria:

(1) Credibility. To ensure the reliability of the coding, members of the research team undertook independent classification of samples of data. The initial analysis and thematic coding of the transcripts was carried out by a research assistant. Two research team members experienced with qualitative analysis independently reviewed a sample of 10% of this thematic coding using the original codebook. The secondary reviewers then met with the primary reviewer to identify agreements and divergences by agreeing on a revised codebook, which subsequently was applied to all transcripts. A second reliability check was performed by two different members of the research team and the primary coder until agreement was attained.
(2) Dependability. To ensure the analysis process was logical, traceable and clearly documented, the processes of categorising and coding were documented to facilitate their consistent application by multiple researchers on the team.

(3) Confirmability. The documentation was enacted to demonstrate how conclusions and interpretations were reached. The codes were grouped into two higher-order categories: (a) framework-relevant codes based on the analysis conducted by Huber et al. (2024); and (b) other codes.

Phase 2. Online survey

Participants and recruitment

A multi-faceted approach was used to recruit participants from the identified groups:

- Email distribution: The survey was emailed to a targeted list of 148 accrediting body representatives, institutional policy-makers, educators and employers, and an additional 61 students from our collective networks, across various higher education institutions.
- Professional networks: The survey link was shared through professional networks and associations related to higher education assessment (e.g. LinkedIn groups, Twitter feeds, conferences, and webinar announcements).
- Snowball sampling: Respondents were encouraged to forward the survey to interested colleagues.
- Reminder emails: Periodic reminder emails were sent to potential participants.
- Incentives: Students were incentivised through a prize draw for one of five vouchers. They could choose to separately provide their name and email address.

A total of 203 people completed the survey: students ($n = 112$; 55.2%) of whom the majority were undergraduates ($n = 88$; 78%); educators ($n = 64$; 31.5%) including course coordinators, lecturers, program coordinators, educational designers, and tutors; institutional policy-makers ($n = 16$; 7.9%); experienced employers of graduates ($n = 9$; 4.4%) most of whom ($n = 5$, 56%) identified as having worked with up to 10 graduates; and accrediting body representatives ($n = 2$; 1%). We acknowledge that the proportions of respondents may not be representative of the underlying populations. Although having responses from two accrediting body representatives is an artefact of our setting, the low number precluded including their survey responses in the analysis. The quantitative analysis was based on 201 respondents, most of whom identified their primary field as Business (38.7%) or one of Accounting or Finance (34%).

Materials

The survey comprised three sections.

1. *Demographic questions*. Participants identified their roles and their main field or discipline.
2. *Importance of framework elements for formative and summative assessment*. We first provided definitions of digital assessment, formative assessment, and summative assessment. We then asked participants to rate the importance of each of (1) learning outcomes; (2) academic integrity; (3) student experience; (4) authenticity; (5) information integrity; (6) quality feedback; and (7) equity of access for each of (a) formative and (b) summative assessments, a total of 14 questions. The scale ranged from 0 'Not important at all' to 4 'Essential', with a further option to indicate that the respondent was unsure or regarded the parameter as not applicable to the assessment type. A final open-ended question asked participants about other key considerations for designing assessments.

Other factors affecting assessment design: genAI and contextual factors

The final section of the survey asked about: (1) the extent to which participants saw the seven framework considerations being impacted by genAI, rated from 0 'Not at all' to 4 'Extremely', with a further option for 'Unsure'; and (2) the importance of considering five broader contextual factors when

designing assessment tasks, using a scale from 1 'Not important at all' to 5 'Essential', with a further option for 'Not Applicable/Unsure'.

We acknowledge several limitations in this study: the results may not be generalisable beyond the time and the context in which we collected data; the stakeholders we engaged may not represent the broader populations from which they are drawn; responses regarding digital assessments may be limited by participants' experience and their exposure to genAI; and our use of ambiguous terms such as *authenticity* may have affected responses in phase 1.

Findings

Phase 1. Focus groups and interviews

Based on our thematic analysis of qualitative data, all five stakeholder groups agreed with the current framework design considerations, although information integrity and equity of access were less prominent issues. Two new elements, 'purpose' and 'technology', were identified as elements missing from the framework.

Our analysis shows that purpose is foundational for assessment design and should be included in the framework. This was emphasised particularly by employers, 'you haven't mentioned anything about the goals' and by educational policy-makers, 'every faculty will have different assessments that feed the learning outcomes and what they're trying to achieve', and educators, who discussed differences in formative and summative assessment design in relation to quality feedback, authenticity, and student experience.

Technology, including genAI, emerges as a missing element that influences assessment design and facilitates its implementation. Educators noted, 'another enabler would be the technology itself and the infrastructure'. Similarly, policy-makers highlighted, 'ensuring upskilling of staff to use online technology in marking and feedback … is a major priority'. This view was echoed by accrediting body representatives, who expressed concerns, 'technical competence [is] missing from the framework … does the assessment address improving the technical competence of students?'

Educators

Educator responses concentrated on student experience, feedback, and authenticity. They highlighted that students experience issues with clarity of assessment instructions. Providing feedback to large cohorts is challenging and educators frequently mentioned new technology, including genAI, that could help provide quality feedback at scale, while recognising the risks around 'lack of awareness of how they can streamline feedback' and the 'integrity' of marking using AI. Authentic assessments and pushing students outside their comfort zones were raised as important to motivate students to engage and learn, do their own work, 'focus on the process' and decrease risk of academic misconduct. Educators emphasised their agency in the assessment design, including issues of requirements, alignment, workload, and sustainability.

Students

Students mainly focused on feedback, student experience, and authenticity. Unlike educators who saw provision of feedback as challenging, particularly at scale, students emphasised the need for detailed feedback, with one reporting 'I've requested feedback, and I was ghosted'. Higher cognitive load and frustration about unclear instructions or little guidance were mentioned, with requests to use 'rubrics and guidelines that would eliminate confusion'. Groupwork was frequently cited as a poor student experience due to concerns of peers' academic misconduct that is time consuming to monitor. Students spoke at length about the authenticity of assessments and reported higher engagement with assessments linked to the future of work because they 'teach us very useful

skills and knowledge', including those that embedded technologies: 'I want to be in the front row seat so that I don't get left behind in the workplace'.

Employers of graduates

Employers of graduates emphasised feedback, authenticity and student experience. While the terminology of authentic assessment was somewhat unclear to the group, they described it as 'tricky', and its understanding 'depends somewhat on the discipline'. This ambiguity provides analytical value because it signals a limit of the shared language between higher education and industry. Employers, who are unlikely to engage with the language of pedagogy, consider authenticity in the application of knowledge, with one respondent recommending authenticity be explicitly defined as 'practical applicability' in relation to 'student learning' while another suggested 'usefulness in an industry context'. Our analysis identifies that the same term can have different meanings to different groups, suggesting a need for open dialogue between employers and universities be maintained to ensure a shared understanding. Employers valued assessments that reflect real-world practices, while recognising the difficulty of connecting assessment performance to workplace outcomes. The importance of student experiences that encompass learning, wellbeing and support were also highlighted. Employers emphasised assessments should promote 'student learning, student development', instead of serving simply as grading mechanisms. They raised questions about how students engaged with feedback and stressed that 'learning as a skill' is crucial in the workplace. Other key skills often overlooked in assessments include: 'soft skills, socialisation, stakeholder engagement … rarely appear in any assessment', stressing the need to promote student autonomy, reflection, and preparation for less familiar assessment formats.

Institutional policy-makers

Policy-makers appeared to adopt a more holistic view of assessment and emphasised two contextual factors – institutional policies and technology – and their role in enhancing other framework elements, particularly student experience and feedback: 'We spend quite a lot of time thinking about how these online assessment tools, online assessment processes support students'. They noted the importance of adapting policies to drive changes in assessment design to enhance authenticity, manage risks, and ensure academic integrity in a world of ongoing technological change, and recognised that institutional policies are influenced by regulatory bodies. Technology emerged as an important element influencing both assessment design and implementation 'assumptions that university students have the most up-to-date technology are sometimes misplaced'. Policy-makers expressed the need to reconsider what is important to assess, emphasising a shift away from assessing outputs to assessing learning processes, while ensuring that staff are well-equipped to navigate this shift. One policy-maker noted the need to 'prioritise learning for students, engaging with the appropriate areas of curriculum and obviously demonstrate the program-based learning outcomes', adding that approach 'can actually benefit the academic'.

Accrediting body representatives

Accrediting bodies highlighted authenticity, academic integrity, and technology. Authenticity was regarded as essential for accrediting degrees to meet real-world needs. As one respondent stated: 'make sure that as that course is being designed, they've actually got some good input from industry'. Representatives suggested having more academics with real-world experience and assessment materials designed by people working in industry, to reduce the academia-industry gap. Some accrediting bodies required that 'particular subjects are taught by people with a particular qualification in industry', however this presents a paradox due to differing academic employment requirements in universities. They suggested that educators should employ alternative assessments to develop and master graduate qualities, 'we really have to try and help them demonstrate the rigour behind [alternative online assessments] and all the rest of it'. Like policy-makers, accrediting body representatives argued that focus should be on the process of learning and higher order skills

to meet emerging industry needs: 'the art is, for me, getting … the right questions so that the student is forced to do their own thinking'. Technology was emphasised as a platform to facilitate assessment and learning, and a risk to academic integrity. They prioritised integrity to ensure that the qualifications accredited are awarded to candidates who have undertaken the assessments: '(universities) had to prove to us that they were upholding the academic integrity and the rigor'.

The findings of the interviews and focus groups guided modifications to the design of the online survey from the work of Huber et al. (2024) to include specific questions about the purpose of assessment to assure course and learning outcomes and a section concerning technology, particularly genAI.

Phase 2. Online survey

Table 2 presents the means and standard deviations of importance ratings for each of the seven design considerations for formative and summative assessments for educators, students, policy-makers, and employers of graduates. From Table 2, for the sample overall, all framework parameters were rated as 'very important' (3) or higher, except for authenticity which was rated slightly below 3 for formative assessments, probably due to the low ratings of importance from both educators and policy-makers who together comprised 40% of the sample completing the survey. In general, these ratings suggest that the framework, originally developed with educators only, is broadly validated by all stakeholders.

Six of the seven parameters were rated as less important for formative compared to summative assessments overall, perhaps reflecting the relative importance of summative assessments compared to formative assessments for final grades, where summative assessments are usually weighted more highly, and the definition provided to respondents stated that summative assessments are 'mainly important for assuring learning outcomes and determining marks'. The exception to this was quality feedback, which was rated as more important for formative assessments, consistent with the definition provided that formative assessments are 'mainly intended for student feedback and learning'. These findings suggest that respondents applied the framework components validly.

ANOVA with post-hoc Bonferroni tests was used to explore whether there were significant differences in the ratings between the groups, with homogeneity of variance assessed using the Levene's test. The only significant difference for formative assessments was for quality feedback which was rated more highly by students than educators ($F(3,196) = 2.95$, $p = .034$). For summative assessments there were significant differences in the importance ratings of purpose ($F(3,194) = 11.62$, $p < .001$), academic integrity ($F(3,197) = 6.77$, $p = .001$), student experience ($F(3,194) = 2.95$, $p = .034$), and quality feedback ($F(3,196) = 3.0$, $p = .032$). Follow-up tests indicated that educators rated purpose and academic integrity as more important for summative assessments than did students, and students rated student experience as more important for summative assessments than did employers. While there was a significant overall difference for quality feedback associated with summative assessments, these were not attributable to specific pairwise differences.

Table 3 shows the mean stakeholder ratings for the importance of considering framework contextual elements in assessment design. The most highly-rated contextual factor across all respondents, and for each stakeholder group, is available resourcing, with professional accrediting body requirements second in overall importance ratings. The one-way ANOVA indicated that there were significant differences between stakeholder groups in the importance ratings for contextual factors for both number of students taking the assessment ('scale'; $F(3,190) = 2.78$, $p = .043$) and institutional policies ($F(3,192) = 7.40$, $p < .001$). Follow-up tests attribute these differences to educators rating the importance of both factors more highly than students (see Table 4).

The final section of the survey asked participants to rate the impact of genAI on the framework dimensions. In the interviews and focus groups, only educators specifically mentioned genAI as a tool to help provide quality feedback at scale, and only students mentioned the additional

Table 2. Mean (standard deviation) ratings of importance of framework elements for formative and summative assessments.

Role	Assessment Type	Purpose	Academic Integrity	Student Experience	Authenticity	Information Integrity	Quality Feedback	Equity of Access
Educators ($n = 64$)	Formative	3.13 (0.95)	3.36 (1.00)	3.02 (0.98)	2.76 (1.04)	3.42 (0.95)	3.77 (0.52)	3.49 (0.89)
	Summative	3.79 (0.45)	3.89 (0.36)	3.17 (0.96)	3.14 (0.88)	3.66 (0.71)	3.17 (0.94)	3.63 (0.88)
Students ($n = 111$)	Formative	3.17 (0.88)	3.21 (1.01)	3.33 (0.82)	3.02 (1.03)	3.22 (1.05)	3.50 (0.78)	3.38 (0.91)
	Summative	3.16 (0.76)	3.45 (0.79)	3.41 (0.80)	3.23 (0.92)	3.32 (1.01)	3.44 (0.78)	3.47 (0.87)
Employers ($n = 9$)	Formative	3.22 (0.97)	3.56 (0.73)	3.56 (0.73)	3.67 (0.50)	3.78 (0.67)	3.56 (0.73)	3.78 (0.44)
	Summative	3.67 (0.71)	3.56 (0.73)	2.56 (1.67)	3.11 (1.36)	3.44 (0.88)	2.89 (1.17)	3.56 (0.73)
Policy-makers ($n = 16$)	Formative	2.93 (1.10)	2.80 (1.37)	3.20 (0.86)	2.87 (1.06)	3.73 (0.59)	3.87 (0.35)	3.47 (0.92)
	Summative	3.47 (0.99)	3.80 (0.41)	3.27 (0.70)	3.20 (0.86)	3.73 (0.59)	3.67 (0.62)	3.80 (0.56)
Total ($n = 200$)	Formative	3.14 (0.92)	3.24 (1.03)	3.23 (0.88)	2.95 (1.03)	3.35 (0.99)	3.62 (0.69)	3.44 (0.89)
	Summative	3.41 (0.75)	3.62 (0.68)	3.28 (0.91)	3.19 (0.92)	3.47 (0.90)	3.35 (0.86)	3.55 (0.85)

Table 3. Mean (standard deviation) ratings of importance of considering contextual factors when designing assessment tasks.

Roles	Scale	Institutional Policies	Technology	Accreditation	Resourcing
Educators ($n = 64$)	3.08 (0.95)	3.56 (0.71)	3.43 (0.71)	3.51 (0.84)	3.69 (0.61)
Students ($n = 111$)	2.65 (1.10)	2.99 (0.95)	3.14 (0.98)	3.18 (0.86)	3.47 (0.73)
Employers ($n = 9$)	2.33 (1.12)	2.78 (1.30)	3.00 (0.87)	2.89 (1.17)	3.67 (0.50)
Policy-makers ($n = 16$)	2.87 (1.19)	3.60 (0.63)	3.53 (0.64)	3.53 (0.64)	3.73 (0.59)
Total Sample ($n = 200$)	**2.79 (1.07)**	**3.21 (0.92)**	**3.26 (0.88)**	**3.30 (0.87)**	**3.57 (0.68)**

Table 4. Mean (standard deviation) ratings of extent to which genAI would impact framework elements.

Role	Purpose	Academic Integrity	Authenticity	Information Integrity	Quality Feedback	Student Experience	Equity of Access
Educators ($n = 64$)	2.47 (1.21)	3.62 (0.72)	3.03 (1.02)	3.08 (1.16)	2.66 (1.16)	2.83 (1.03)	2.79 (1.16)
Students ($n = 111$)	2.25 (1.07)	2.96 (1.09)	2.83 (1.08)	2.63 (1.09)	2.25 (1.20)	2.44 (1.06)	2.30 (1.22)
Employers ($n = 9$)	2.33 (0.87)	3.11 (1.27)	2.89 (1.27)	2.78 (1.09)	2.67 (1.12)	2.63 (0.52)	2.67 (0.87)
Policy-makers ($n = 16$)	2.53 (0.83)	3.40 (0.74)	3.00 (0.76)	2.93 (0.92)	2.47 (1.36)	2.47 (1.13)	2.50 (1.34)
Total ($n = 200$)	**2.35 (1.09)**	**3.22 (1.01)**	**2.91 (1.05)**	**2.81 (1.11)**	**2.42 (1.20)**	**2.57 (1.05)**	**2.49 (1.21)**

incentive to engage when genAI is included in assessment tasks, due to its importance for their future work. From the survey, the highest rating for the impact of genAI was for academic integrity (see Table 4) and the one-way ANOVA indicated that there were significant differences between stakeholder groups for academic integrity only ($F(3,194) = 6.38$, $p < .001$). Again, follow-up tests attribute this to educators rating the impact of genAI on academic integrity more highly than students.

Discussion: a revised framework and future directions

The aim of this study was to understand the perspectives and priorities of key stakeholders concerned with achieving quality assessment in business education (Phase 1) and to evaluate the alignment of stakeholder priorities with Huber et al. (2024) framework, thereby validating and modifying the framework (Phase 2). Whilst our findings affirm the multidimensional nature of assessment design, they also reveal the tensions and assumptions shaping stakeholder priorities and contextual constraints. The findings do more than map variation, they indicate key challenges in aligning assessment design with values such as integrity, authenticity and equity, all of which are under increasing pressure in an evolving digital landscape. We first discuss affirmed framework design considerations before addressing newly identified elements.

Educators commented on practical issues in assessments and highlighted the need to identify ways in which integrity and authenticity can be balanced and enhanced. Students and accrediting bodies were concerned with authenticity so that assessments prepare students for work. Students also emphasised practical experiences in completing assessments, including the need for feedback to enhance learning. Employers prioritised both the outcome and the process of assessment, reflecting their interest in graduate skills and attributes, including the ability to learn from their experiences. Policy-makers and accrediting bodies valued integrity and echoed employers' interest in the process of learning. Policy-makers and educators shared an interest in feedback efficiency at scale.

Stakeholders were relatively silent on information integrity and equity of access in Phase 1, perhaps because these are assumed to be fundamental premises from which assessment design begins. Alternatively, this silence may signify a deeper misalignment between policy rhetoric and pedagogical realities that obscures and silences the structural power asymmetries embedded in digital and AI-mediated assessment environments. If we accept that learning begins with experience, and that experience is socially constructed, then learning in higher education must reflect the experiences of students and educators (Köseoğlu, Veletsianos, and Rowell 2023). From a critical digital

pedagogy perspective, focus group and interview data can shine much needed light on the unspoken pressures that impact stakeholders in higher education globally, heavily amplified by the rapid transitions driven by genAI. By including students, educators, policy-makers, accrediting bodies and employers it is possible to begin to identify the beneficiaries of the shift to digital and AI-driven assessment, those who bear its risks, and those whose voices determine how such systems are configured (Fawns et al. 2025; Rapanta et al. 2025). In the current Australian higher education context, questions of agency, justice, and voice are central, and neglecting these dynamics risks reinforcing existing digital inequalities and exclusion, undermining fairness and credibility, and exposing students to unauthorised uses of personal data as online and AI-driven assessment proliferates. The survey findings showed that educators rated purpose and academic integrity for summative assessments (Table 2) and both scale of assessment and institutional policies (Table 3) higher than students, reflecting the dominant role of educators as gatekeepers of academic standards. Working within frameworks of institutional policies and changes to funding models has resulted in pressure on resources, where recent reforms in Australian higher education funding and increased corporatisation have placed growing administrative burdens on academic staff, contributing to heavier workloads and reduced time available for core activities such as teaching and research (Woelert et al. 2025).

The educators surveyed are concerned about genAI's impact on academic integrity, signalling a reassertion of disciplinary control in response to technological uncertainty. Students are less concerned about this, perhaps reflecting a more instrumental and future-oriented view that recognises the extent to which genAI has already transformed the workplaces they will enter. Findings such as these showing differing stakeholder perspectives assist educators to take a 'big picture' view of assessment design that allows for 'trade-offs' within courses, balancing the outcomes sought by stakeholder groups.

Across Phases 1 and 2, the findings are informative about the priorities of stakeholder groups and their perspectives on the framework parameters, and suggest that while stakeholders may prioritise similar parameters (Phase 2), their reasons may differ (Phase 1). For example, the data presented in Table 2 indicates that all groups rate authenticity as very important for summative assessments, with average ratings in a narrow range between 3.11 (employers) and 3.23 (students). However, the perspectives on why authenticity is important vary. Educators viewed authenticity as a pedagogical strategy to enhance student engagement, deepen learning and decrease the risk of academic misconduct, while students understood authenticity primarily through a pragmatic lens, as a means of preparing for the workplace and demonstrating employability skills. Accrediting bodies linked authenticity to the alignment between assessments and professional standards, while employers, though equally supportive of authentic learning, expressed uncertainty about the term itself. Some found the term 'authenticity' ambiguous and confusing, offering alternative terms to define this construct, and preferring language that emphasised relevance, applicability and transferability of learning to real-world practice.

This ambiguity is a significant finding in itself, because it highlights how the language of authenticity, commonly used in educational discourse, can obscure rather than clarify shared expectations of quality assessment. This interpretive gap, for instance, between how educators and industry view authenticity, exposes the challenge of developing assessments that are simultaneously pedagogically sound, professionally relevant, and conceptually transparent across diverse stakeholder groups. As a result, in Phase 2, we took care to define the terms used.

The findings also highlight the socially constructed nature of assessment language. Assessment discourses often rely on static, techno-rational models that fail to account for the interpretive and contextual dimensions of meaning-making (Gonsalves and Lin 2025). Our data suggest that clarity of terminology, especially in multi-stakeholder environments, is a matter of communication and epistemic alignment. Without shared understanding, the goal of 'authentic assessment' risks fragmentation: what educators design as authentic may not be recognised as such by employers, students, or accrediting bodies.

We argue that authenticity in assessment design should be understood as a relational construct negotiated among stakeholders. This aligns with work showing authenticity as subjective, situated, and shaped by users' perceptions and contexts (Gulikers, Bastiaens, and Kirschner 2004). Recent scholarship has broadened the concept, noting that authenticity can take different forms that emphasise engagement, professional relevance, or participation in meaningful practices (Ajjawi et al. 2024; Fawns et al. 2025). Our findings reflect these differences, as stakeholders endorsed authenticity while interpreting it in distinct ways. Recognising authenticity in its multiplicity invites educators and policymakers to co-define what counts as credible evidence of learning, supporting more coherent assessment design.

Students consistently expressed the desire for detailed feedback, however, educators reported challenges with providing feedback, particularly in scaling its delivery in large classes (see also Cram et al. 2022). Employers saw learning from feedback as a critical workplace skill, reinforcing the need for students to practice it during their education. Feedback must go beyond reflecting on past performance and serve as a feed-forward mechanism to guide students in future assessments (Murphy and Barry 2016). Ensuring that feedback is actionable and relevant to subsequent tasks enhances learning and fosters continuous development, an essential quality in academic and professional contexts (Winstone and Boud 2022). These divergent concerns echo longstanding tensions between pedagogical care and institutional performance metrics. As feedback becomes increasingly automated, the challenge is not only technical but epistemological: how to ensure feedback fosters dialogic learning rather than compliance.

The contextual factor prioritised by each group reflects their priorities. Policy-makers referred to policies, accrediting bodies highlighted accreditation, while educators and students referred to training on assessments. Educators highlighted the need for sustainable resourcing and practices in assessment design. One described 'sustainable assessment practices … set up in a way that is easily changed each semester'. Given the future challenges regarding funding in higher education, finding creative solutions to resourcing problems will be critical to success (Jackson 2024).

Our data identified two new factors that impact the quality of digital assessment design: purpose and technology (see Table 5). Purpose is the taken-for-granted foundation underpinning quality assessment design, which was identified for explicit inclusion by the wider group of stakeholders consulted in the current study. Purpose has two facets: consideration of the intended learning outcomes, and the broader rationale(s) of assessments. Survey respondents distinguished between summative and formative assessments that varied in purpose: in Phase 2, all stakeholder groups rated quality feedback highest for formative assessments and academic integrity highest for summative assessments. Employers, policy-makers and accrediting bodies also advocated for assessment that developed student capabilities with relevance beyond the course, including the process of learning, the development of soft skills and higher order thinking. In contrast, educators and students focussed more on the immediate value of assessment for learning and grading within the course. Taken together, these perspectives echo Boud and Soler's (2016) notion of sustainable assessment which advocates assessment that facilitates grading as well as both current and future student learning.

Table 5. Additional framework elements.

	Description	Related literature
New design consideration		
Purpose	Rationale and intention underpinning assessment design, incorporating formative and summative aspects, enablement of sustainable learning, and intended learning outcomes.	Boud and Soler 2016; Bearman, Nieminen, and Ajjawi 2023; Nieminen, Bearman, and Ajjawi 2023.
New contextual factor		
Technology	A dynamic sociotechnical ecosystem of digital platforms, including genAI, that integrates into and transforms assessment design, delivery, and management, shaping diverse aspects of learning, teaching and educational processes.	Bearman, Nieminen, and Ajjawi 2023; Fawns et al. 2025.

Technology was highlighted by all stakeholder groups, reflecting the influence of the changing technological environment since Huber et al. (2024). Technology is the dynamic ecosystem of platforms, including genAI, that shapes assessment design, delivery, and management, and influences diverse aspects of learning, teaching and educational processes (Bearman, Nieminen, and Ajjawi 2023; Fawns et al. 2025). We recognise that technology will continue to change, causing future disruptions to assessment as new tools and solutions emerge (Bennett et al. 2017), and requiring ongoing adjustments to assessment design. By including technology as a dynamic, co-constitutive force in assessment design, shaping what is possible, valued, and assessable (Bearman, Nieminen, and Ajjawi 2023), our framework facilitates this process of adjustment and enables illumination of the impact of disruptions on factors such as assessment equity, integrity and authenticity.

While there were differences in stakeholder emphasis during the focus groups, there was consensus in the survey responses that all framework elements were relevant and important to consider for digital assessment design. This includes the ten original design dimensions and contextual factors (Table 1) and the two new elements (Table 5) added to the survey as a result of the attention given them in the focus groups: (a) Technology, previously considered as part of Resourcing, is now included as a contextual factor, and (b) Purpose, an underlying driver of assessment design, is now placed at the centre of other design considerations.

The interplay between the seven design considerations and the five contextual factors illustrated in Figure 1 demonstrates the complexity of meeting diverse stakeholder expectations regarding assessment design. This revised framework identifies key considerations for quality assessment design while facilitating dialogue about contrasting perspectives on the relative importance and interpretation of assessment quality. The findings highlight the value of co-design in assessment design, as collaborative input from diverse and engaged stakeholders can foster shared

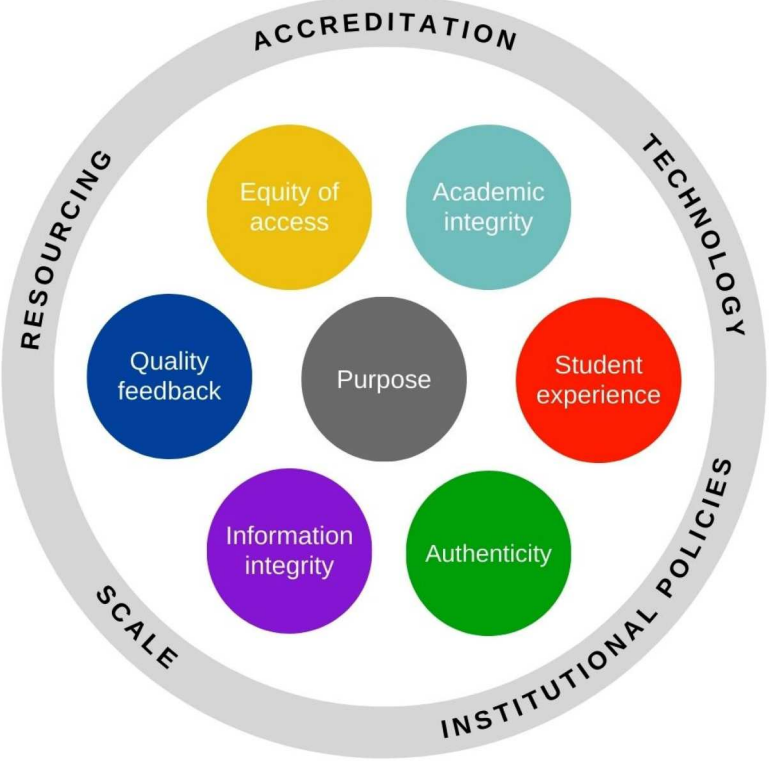

Figure 1. Framework for supporting the design and evaluation of digital assessments.

understanding and accountability. The explicit addition of purpose and technology to the framework responds to the importance stakeholders placed on these factors. While both purpose and technology were implicit in the original framework, their addition as explicit elements enhances the relevance of the framework across the stakeholder groups and encourages alignment with educational and workplace needs.

Understanding the priorities and perspectives of diverse stakeholder groups concerning assessment design is essential to ensure that quality assessment meets stakeholder needs and prepares students for the uncertain world and workplaces that they will enter as graduates and lifelong learners. This study reveals aspects of assessment quality in which key stakeholders are in close alignment, while highlighting factors such as authenticity, purpose and the ongoing impacts of genAI in which contrasting perspectives are substantive and epistemic alignment should not be assumed.

In increasingly disrupted higher education, the framework offers a practical tool for eliciting constructive discussions among diverse assessment stakeholders. It supports assessment co-design within courses and programs, and contributes to broader sector-level understandings of changing assessment practices and impacts of genAI (see e.g. Kizilcec et al. 2024). Engagement across stakeholder groups fosters shared accountability and reveals blind spots that single-group perspectives may miss. Co-design must grapple with power asymmetries, particularly between institutional actors and learners, and between policy imperatives and pedagogical realities. Future research should examine how the framework is adopted and adapted in different settings, and how multi-stakeholder engagement shapes the design of digital assessments.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the University of Sydney Business School's education research funding.

ORCID

Sandris Zeivots http://orcid.org/0000-0001-5450-2530
Sue Wright http://orcid.org/0000-0001-8634-1428
Lynne Harris http://orcid.org/0000-0001-8925-1938
Andrew Cram http://orcid.org/0000-0002-4125-386X
Elaine Huber http://orcid.org/0000-0001-7964-9096
Corina Raduescu http://orcid.org/0000-0002-6281-9842
Amanda White http://orcid.org/0000-0003-1164-7806

References

Ajjawi, R., J. Tai, M. Dollinger, P. Dawson, D. Boud, and M. Bearman. 2024. "From Authentic Assessment to Authenticity in Assessment: Broadening Perspectives." *Assessment & Evaluation in Higher Education* 49 (4): 499–510. https://doi.org/10.1080/02602938.2023.2271193.

Ali, I., A. K. Narayan, and D. Gedera. 2022. "Transforming Assessment in Accounting Education to Align with Online Learning." *Pacific Accounting Review* 34 (4): 536–547. https://doi.org/10.1108/PAR-05-2021-0058.

Bearman, M., P. Dawson, D. Boud, S. Bennett, M. Hall, and E. Molloy. 2016. "Support for Assessment Practice: Developing the Assessment Design Decisions Framework." *Teaching in Higher Education* 21 (5): 545–556. https://doi.org/10.1080/13562517.2016.1160217.

Bearman, M., J. H. Nieminen, and R. Ajjawi. 2023. "Designing Assessment in a Digital World: An Organising Framework." *Assessment & Evaluation in Higher Education* 48 (3): 291–304. https://doi.org/10.1080/02602938.2022.2069674.

Bennett, S., P. Dawson, M. Bearman, E. Molloy, and D. Boud. 2017. "How Technology Shapes Assessment Design: Findings from a Study of University Teachers." *British Journal of Educational Technology* 48 (2): 672–682. https://doi.org/10.1111/bjet.12439.

Bhullar, P. S., M. Joshi, and R. Chugh. 2024. "ChatGPT in Higher Education - A Synthesis of the Literature and a Future Research Agenda." *Education and Information Technologies* 29 (16): 21501–21522. https://doi.org/10.1007/s10639-024-12723-x.

Birks, M., J. Mills, S. Allen, and S. Tee. 2020. "Managing the Mutations: Academic Misconduct in Australia, New Zealand and the UK." *International Journal for Educational Integrity* 16 (1): 1–15. https://doi.org/10.1007/s40979-020-00055-5.

Boud, D., and R. Soler. 2016. "Sustainable Assessment Revisited." *Assessment & Evaluation in Higher Education* 41 (3): 400–413. https://doi.org/10.1080/02602938.2015.1018133.

Butler-Henderson, K., and J. Crawford. 2020. "A Systematic Review of Online Examinations: A Pedagogical Innovation for Scalable Authentication and Integrity." *Computers & Education* 159:104024. https://doi.org/10.1016/j.compedu.2020.104024.

Cavalcanti, A. P., A. Barbosa, R. Carvalho, F. Freitas, Y. S. Tsai, D. Gašević, and R. F. Mello. 2021. "Automatic Feedback in Online Learning Environments: A Systematic Literature Review." *Computers and Education: Artificial Intelligence* 2:100027. https://doi.org/10.1016/j.caeai.2021.100027.

Chan, C. K. Y. 2023. "A Comprehensive AI Policy Education Framework for University Teaching and Learning." *International Journal of Educational Technology in Higher Education* 20 (1): 38. https://doi.org/10.1186/s41239-023-00408-3.

Corbin, T., P. Dawson, K. Nicola-Richmond, and H. Partridge. 2025. "'Where's the Line? It's an Absurd Line': Towards a Framework for Acceptable Uses of AI in Assessment." *Assessment & Evaluation in Higher Education* 50 (5): 705–717. https://doi.org/10.1080/02602938.2025.2456207.

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 and Teaching International* 61 (2): 228–239. https://doi.org/10.1080/14703297.2023.2190148.

Cram, A., L. Harris, C. Raduescu, E. Huber, S. Zeivots, A. Brodzeli, S. Wright, and A. White. 2022. "Online Assessment in Australian University Business Schools: A Snapshot of Usage and Challenges." *ASCILITE Publications* e22181-1–e22181-10. https://doi.org/10.14742/apubs.2022.181.

Cramp, J., J. Medlin, P. Lake, and C. Sharp. 2019. "Lessons Learned from Implementing Remotely Invigilated Online Exams." *Journal of University Teaching and Learning Practice* 16 (1): 1–20. https://doi.org/10.53761/1.16.1.10.

Crawford, J., M. Cowling, and K. Allen. 2023. "Leadership Is Needed for Ethical ChatGPT: Character, Assessment, and Learning Using Artificial Intelligence (AI)." *Journal of University Teaching and Learning Practice* 20 (3): 1–19. https://doi.org/10.53761/1.20.3.02.

Dawson, P., and M. Henderson. 2017. "How Does Technology Enable Scaling up Assessment for Learning?." In *Scaling Up Assessment for Learning in Higher Education*, edited by D. Carless, S. M. Bridges, C. K. Y. Chan, and R. Glofcheski, 209–222. Singapore: Springer. https://doi.org/10.1007/978-981-10-3045-1_14.

Farazouli, A., T. Cerrato Pargman, K. Bolander Laksov, and C. McGrath. 2025. "Navigating Uncertainty: University Teachers' Experiences and Perceptions of Generative Artificial Intelligence in Teaching and Learning." *Studies in Higher Education*, 1–16. https://doi.org/10.1080/03075079.2025.2550766.

Fawns, T., M. Bearman, P. Dawson, J. H. Nieminen, K. Ashford-Rowe, K. Willey, L. X. Jensen, C. Damşa, and N. Press. 2025. "Authentic Assessment: From Panacea to Criticality." *Assessment & Evaluation in Higher Education* 50 (3): 396–408. https://doi.org/10.1080/02602938.2024.2404634.

Gonsalves, C., and Z. Lin. 2025. "Clear in Advance to Whom? Exploring 'Transparency' of Assessment Practices in UK Higher Education Institution Assessment Policy." *Studies in Higher Education* 50 (7): 1454–1470. https://doi.org/10.1080/03075079.2024.2381124.

Grosseck, G., S. Sava, G. Ion, and L. Malita. 2024. *Digital Assessment in Higher Education: Navigating and Researching Challenges and Opportunities*. Singapore: Springer.

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 (3): 67–86. https://doi.org/10.1007/BF02504676.

Hancock, P., J. Birt, P. De Lange, C. Fowler, M. Kavanagh, L. Mitrione, M. Rankin, G. Slaughter, and A. Williams. 2023. "Integrity of Assessments in Challenging Times." *Accounting Education* 32 (5): 501–522. https://doi.org/10.1080/09639284.2022.2137818.

Heil, J., and D. Ifenthaler. 2023. "Online assessment in higher education: A systematic review." *Online Learning* 27 (1): 187–218. https://doi.org/10.24059/olj.v27i1.3398.

Henderson, M., M. Bearman, J. Chung, T. Fawns, S. Buckingham Shum, K. E. Matthews, and J. de Mello Heredia. 2025. "Comparing Generative AI and Teacher Feedback: Student Perceptions of Usefulness and Trustworthiness." *Assessment & Evaluation in Higher Education*, 1–16. https://doi.org/10.1080/02602938.2025.2502582.

Hopfenbeck, T. N., Z. Zhang, S. Z. Sun, P. Robertson, and J. A. McGrane. 2023. "Challenges and Opportunities for Classroom-Based Formative Assessment and AI: A Perspective Article." *Frontiers in Education* 8:1270700. https://doi.org/10.3389/feduc.2023.1270700.

Huber, E., L. Harris, S. Wright, A. White, C. Raduescu, S. Zeivots, A. Cram, and A. Brodzeli. 2024. "Towards a Framework for Designing and Evaluating Online Assessments in Business Education." *Assessment & Evaluation in Higher Education* 49 (1): 102–116. https://doi.org/10.1080/02602938.2023.2183487.

Jackson, C. 2024, October 14. "Unis Heading for a Funding Cliff under Policy Chaos." Universities Australia. https://universitiesaustralia.edu.au/media-item/unis-heading-for-a-funding-cliff-under-policy-chaos.

Kharbat, F. F., and A. S. Abu Daabes. 2021. "E-proctored Exams During the COVID-19 Pandemic: A Close Understanding." *Education and Information Technologies* 26 (6): 6589–6605. https://doi.org/10.1007/s10639-021-10458-7.

Kizilcec, R. F, E. Huber, E. C. Papanastasiou, A. Cram, C. A. Makridis, A. Smolansky, S. Zeivots, and C. Raduescu. 2024. "Perceived Impact of Generative AI on Assessments: Comparing Educator and Student Perspectives in Australia, Cyprus, and the United States." *Computers and Education: Artificial Intelligence* 7: 100269. https://doi.org/10.1016/j.caeai.2024.100269.

Köseoðlu, S., G. Veletsianos, and C. Rowell2023. *Critical Digital Pedagogy in Higher Education*. Athabasca: Athabasca University Press.

Lodge, J. M., S. Howard, M. Bearman, P. Dawson, and S. Agostinho. 2023. *Assessment Reform for the age of Artificial Intelligence*. Melbourne: Australian Government, Tertiary Education Quality and Standards Agency.

Martel, J. S. S., and A. Pérez Garcias. 2024. "A Model of Assessment co-creation in Technology-Enhanced Learning Environments in Higher Education." *Digital Education Review* 45:204–213. https://doi.org/10.1344/der.2024.45.204-213.

Mate, K., and J. Weidenhofer. 2022. "Considerations and Strategies for Effective Online Assessment with a Focus on the Biomedical Sciences." *FASEB BioAdvances* 4 (1): 9–21. https://doi.org/10.1096/fba.2021-00075.

Moorhouse, B. L., M. A. Yeo, and Y. Wan. 2023. "Generative AI Tools and Assessment: Guidelines of the World's Top-Ranking Universities." *Computers and Education Open* 5:100151. https://doi.org/10.1016/j.caeo.2023.100151.

Murphy, K., and S. Barry. 2016. "Feed-Forward: Students Gaining More from Assessment via Deeper Engagement in Video-Recorded Presentations." *Assessment & Evaluation in Higher Education* 41 (2): 213–227. https://doi.org/10.1080/02602938.2014.996206.

Nieminen, J. H., M. Bearman, and R. Ajjawi. 2023. "Designing the Digital in Authentic Assessment: Is It fit for Purpose?" *Assessment & Evaluation in Higher Education* 48 (4): 529–543. https://doi.org/10.1080/02602938.2022.2089627.

Nowell, L. S., J. M. Norris, D. E. White, and N. J. Moules. 2017. "Thematic Analysis: Striving to Meet the Trustworthiness Criteria." *International Journal of Qualitative Methods* 16 (1): 1609406917733847. https://doi.org/10.1177/1609406917733847.

Okada, A., I. Noguera, L. Alexieva, A. Rozeva, S. Kocdar, F. Brouns, T. Ladonlahti, D. Whitelock, and A. E. Guerrero-Roldán. 2019. "Pedagogical Approaches for e-Assessment with Authentication and Authorship Verification in Higher Education." *British Journal of Educational Technology* 50 (6): 3264–3282. https://doi.org/10.1111/bjet.12733.

Perkins, M., J. Roe, and L. Furze. 2025. "Reimagining the Artificial Intelligence Assessment Scale: A Refined Framework for Educational Assessment." *Journal of University Teaching and Learning Practice* 22. https://doi.org/10.53761/rrm4y757.

Rapanta, C., I. Bhatt, A. Bozkurt, L. A. Chubb, C. Erb, I. Forsler, K. Gravett, et al. 2025. "Critical GenAI Literacy: Postdigital Configurations." *Postdigital Science and Education* 7 (4): 1296–1333. https://doi.org/10.1007/s42438-025-00573-w.

Salinas-Navarro, D. E., E. Vilalta-Perdomo, R. Michel-Villarreal, and L. Montesinos. 2024. "Using Generative Artificial Intelligence Tools to Explain and Enhance Experiential Learning for Authentic Assessment." *Education Sciences* 14 (1): 83. https://doi.org/10.3390/educsci14010083.

Sotiriadou, P., D. Logan, A. Daly, and R. Guest. 2020. "The Role of Authentic Assessment to Preserve Academic Integrity and Promote Skill Development and Employability." *Studies in Higher Education* 45 (11): 2132–2148. https://doi.org/10.1080/03075079.2019.1582015.

Stephenson, B., and A. Harvey. 2022. "Student Equity in the age of AI-Enabled Assessment: Towards a Politics of Inclusion." In *Assessment for Inclusion in Higher Education*, edited by R. Ajjawi, J. Tai, D. Boud, and T. J. de St Jorre, 120–130. Abingdon: Routledge.

Swiecki, Z., H. Khosravi, G. Chen, R. Martinez-Maldonado, J. M. Lodge, S. Milligan, N. Selwyn, and D. Gašević. 2022. "Assessment in the Age of Artificial Intelligence." *Computers and Education: Artificial Intelligence* 3:100075. https://doi.org/10.1016/j.caeai.2022.100075.

Tai, J., P. Mahoney, R. Ajjawi, M. Bearman, J. Dargusch, M. Dracup, and L. Harris. 2023. "How Are Examinations Inclusive for Students with Disabilities in Higher Education? A Sociomaterial Analysis." *Assessment & Evaluation in Higher Education* 48 (3): 390–402. https://doi.org/10.1080/02602938.2022.2077910.

TEQSA. 2024. "Artificial Intelligence." Tertiary Education Quality and Standards Agency. https://www.teqsa.gov.au/guides-resources/higher-education-good-practice-hub/artificial-intelligence.

Timmis, S., P. Broadfoot, R. Sutherland, and A. Oldfield. 2016. "Rethinking Assessment in a Digital age: Opportunities, Challenges and Risks." *British Educational Research Journal* 42 (3): 454–476. https://doi.org/10.1002/berj.3215.

Viberg, O., C. Mutimukwe, S. Hrastinski, T. Cerratto-Pargman, and J. Lilliesköld. 2024. "Exploring Teachers' (Future) Digital Assessment Practices in Higher Education: Instrument and Model Development." *British Journal of Educational Technology* 55:2597–2616. https://doi.org/10.1111/bjet.13462.

Walton, J., M. Bearman, N. Crawford, J. Tai, and D. Boud. 2025. "How University Students Work on Assessment Tasks with Generative Artificial Intelligence: Matters of Judgement." *Assessment & Evaluation in Higher Education*, 1–17. https://doi.org/10.1080/02602938.2025.2570328.

Winstone, N. E., and D. Boud. 2022. "The Need to Disentangle Assessment and Feedback in Higher Education." *Studies in Higher Education* 47 (3): 656–667. https://doi.org/10.1080/03075079.2020.1779687.

Woelert, P., J. Chesters, M. Martinussen, and J. Gannaway. 2025. "Administrative Burden in Australian Universities: Insights into Dimensions and Drivers from a Nationwide Survey." *Science and Public Policy*, scaf029. https://doi.org/10.1093/scipol/scaf029.

Zeivots, S., A. Casey, T. Winchester, J. Webster, X. Wang, L. Tan, W. Smeenk, et al. 2025. "Reshaping Higher Education Designs and Futures: Postdigital Co-Design with Generative Artificial Intelligence." *Postdigital Science and Education* 7:1334–1374. https://doi.org/10.1007/s42438-025-00595-4.

Zeivots, S., N. Hopwood, D. Wardak, and A. Cram. 2025. "Co-design Practice in Higher Education: Practice Theory Insights into Collaborative Curriculum Development." *Higher Education Research & Development* 44 (3): 769–783. https://doi.org/10.1080/07294360.2024.2410269.