

Authentic Assessment in the Age of Generative Artificial Intelligence: Pedagogical Innovation to Institutional Infrastructure

Sharon Lehane¹ , Angela Wright² , Pio Fenton³ 

¹Department of Management and Enterprise, Munster Technological University, Ireland, ²Head of Teaching and Learning Unit, AnSEO, Student Engagement and Arts Office, Munster Technological University, Ireland, ³Vice President People and Culture, Munster Technological University, Ireland.

How to cite: Lehane, S.; Wright, A.; Fenton, P. (2026). Authentic Assessment in the Age of Generative Artificial Intelligence: Pedagogical Innovation to Institutional Infrastructure. In: 12th International Conference on Higher Education Advances (HEAd'26). Valencia, 15-18 June 2026. <https://doi.org/10.4995/HEAd26.2026.21804>

Abstract

The rapid emergence of generative artificial intelligence (GenAI) constitutes a persistent “wicked” problem for higher education, threatening established assumptions concerning authorship, assessment validity, and learning assurance. Institutional responses have prioritised procedural compliance, however, this approach is proving increasingly inadequate. Although authentic assessment (AA) has been widely promoted as a means of enhancing learning relevance, engagement, and employability, its implementation remains highly dependent on individual academic initiative and informal experimentation.

This study employs reflexive thematic analysis to identify the institutional conditions required for the sustainable implementation of AA in GenAI-enabled environments. Findings reveal a pronounced paradox of practice; while 84% of survey respondents report using AA, 49% have received no formal training, and 61% identify workload as their primary barrier. Experts further highlight systemic constraints relating to technological ambiguity and policy fragmentation. In response, this paper proposes a systemic framework that positions AA as organisational infrastructure rather than discretionary practice.

Keywords: *authentic assessment; generative artificial intelligence; assessment governance; academic integrity; institutional capacity; university assessment framework*

1. Introduction

Assessment in higher education is currently confronting a profound governance challenge. The rapid development of GenAI has rendered traditional assessment increasingly unfit for purpose.

Tasks that previously functioned as reliable indicators of student learning can now be completed using GenAI, which complicates processes of educator judgement, assessment verification, and institutional certification (Corbin *et al.*, 2025). As a result, institutions are increasingly required to reconsider, not only assessment formats, but also the organisational systems that assure academic standards. Institutional responses to this disruption have, to date, largely prioritised procedural interventions, including policy revisions and academic integrity statements. While such measures provide clear signals of regulatory compliance, they have proven insufficient in contexts characterised by technological volatility, disciplinary variation, and professional uncertainty. Detection systems remain unreliable, policy guidance is frequently ambiguous, and responsibility for managing risk is often devolved to individual lecturers. Consequently, many institutions continue to rely on fragmented, reactive governance arrangements.

2. Literature and Context

Recent scholarship conceptualises GenAI and assessment as a wicked problem: a persistent institutional condition marked by uncertainty, competing values, and the absence of definitive solutions (Corbin *et al.*, 2025; Lodge *et al.*, 2025). Within this context, AA has been widely promoted as a pedagogical response. Through the emphasis of contextualised judgement, applied problem-solving, and professional relevance, AA is widely understood to enhance student engagement, employability development and to support “integrity by design” (Lehane *et al.*, 2024). Empirical studies demonstrate that AA can enhance learning validation and strengthen students’ evaluation judgement and sense of ownership over academic work (Lehane *et al.*, 2025). Despite its recognised benefits, however, AA implementation remains fragmented and highly dependent on individual academic initiative. Much of the existing literature presents successful AA practices as transferable exemplars without adequate attention being given to institutional conditions. For this reason, AA frequently operates as discretionary practice sustained through informal networks and personal commitment. This reliance on “heroic” individualism renders AA innovation vulnerable to workload pressures, policy misalignment, and professional burnout (Gunder *et al.*, 2025; Nicola-Richmond *et al.*, 2025). Corbin *et al.* (2025) argue that governing wicked problems requires the development of “institutional permissions”: formal and informal authorisations that enable professionals to compromise, experiment, and revise pedagogical and assessment practices without fear of sanction. In the absence of such permissions, staff are constrained by compliance-oriented systems that discourage experimentation and reflective, iterative redesign. From this perspective, GenAI does not merely intensify existing integrity risks but exposes fundamental limitations in rule-based governance. Effective institutional responses, therefore, depend less on definitive solutions than on the cultivation of organisational capacities for ongoing negotiation, iteration and learning (Lodge *et al.*, 2025). Lodge *et al.* (2025) introduce the concept of “fjord modules” to describe isolated pockets of pedagogical excellence embedded within fragmented curricular

landscapes. These modules often showcase innovative AA practices, strong industry engagement, and high student satisfaction; however, they remain disconnected from broader programme trajectories. Fjord modules exemplify the limits of bottom-up reform; while they demonstrate what is pedagogically possible, their isolation limits their capacity to influence institutional practice, as local lecturer experimentation is rarely systemised, shared, or embedded within organisational structures (Gunder *et al.*, 2025). This fjord concept, therefore, provides a powerful theoretical justification for systemic approaches to assessment governance. Without coordinated infrastructures for curriculum mapping, professional learning, and evidence sharing, successful AA practices fail to generate cumulative institutional benefit.

This paper constitutes the final phase of a three-part study examining AA in contemporary higher education. The first study focused on pedagogical design; a framework for integrity-oriented AA was developed by examining how scaffolded tasks, exemplar use, and reflective components could support learning assurance and reduce misconduct (Lehane *et al.*, 2024). The second study explored student experiences of AA and demonstrated that students perceived AA approaches as more meaningful, engaging, and professionally relevant due to stronger ownership of learning and ethical engagement (Lehane *et al.*, 2025). While these studies established the pedagogical value of AA, they also revealed a persistent structural tension: successful practices were sustained largely through individual effort rather than institutional systems, thus raising critical questions in relation to AA scalability. The present study, therefore, shifts analytical attention from individual classrooms to organisational capacity, by examining how institutions govern, resource, and encourage AA under conditions of technological disruption. An institutional framework is proposed in this study to transition AA from isolated experiments into sustainable, university-level infrastructure. The framework integrates pedagogical design, professional expertise, governance alignment, and evidence-informed review within a self-reinforcing architecture. The framework contributes to debates on assessment governance and GenAI disruption by demonstrating how universities can move from reactive compliance toward adaptive institutional stewardship. The framework offers practical guidance for leaders and academics seeking to embed integrity, validity, and learning assurance within coordinated organisational systems. Navigating GenAI issues requires a transition from isolated design to a nested infrastructure in which governance and professional capability are mutually reinforcing. Against this backdrop, the paper asks: *How do academic staff and assessment experts conceptualise authentic assessment, and what institutional capacities are required to govern its implementation in GenAI-enabled higher education?*

3. Methodology

This study adopted a qualitative interpretivist design to examine how academic staff and assessment experts construct meaning around assessment reform in contexts of disruption. An interpretivist approach is appropriate as this study explores complex educational phenomena

characterised by professional judgement, institutional culture, and context-dependent decisions. This study was designed as a bounded qualitative case study centred on the Faculty of Business at Munster Technological University (MTU), Cork, Ireland, where the faculty was conceptualised as a “village” within a larger institutional ecosystem (Nicola-Richmond *et al.*, 2025). Staff members function as “villagers” navigating day-to-day pedagogical and administrative work, policy ambiguity, and resource constraints, while expert participants represent system-level actors responsible for mediating between regulatory requirements, institutional strategy, and local practice. This triangulation enabled examination of how institutional policies are interpreted, enacted, and adapted locally. Data was collected between June and October 2025 through; [1] An anonymous online survey of academic staff in the Faculty of Business at MTU ($n = 60$), and [2] Semi-structured interviews with assessment experts from each Irish Technological University ($n = 5$). Academic staff participants were recruited through an email distributed to all Faculty of Business lecturers in June 2025, and participation was voluntary and anonymous. Assessment experts were recruited through purposive sampling based on their senior institutional roles and recognised expertise in assessment leadership, academic development, and quality assurance within the Irish Technological University sector. Data was analysed using reflexive thematic analysis. Inductive coding was complemented by deductive engagement with governance and integrity literature. Themes were developed through iterative comparison and negotiation between all researchers.

4. Results

This section presents the findings from the academic staff survey ($n=60$) and semi-structured interviews with national assessment experts ($n=5$). Together, these datasets highlight what may be characterised as a “paradox of practice”: AA has achieved widespread pedagogical legitimacy and adoption yet remains structurally fragile due to limited institutional investment in professional capability, workload recognition, and governance coordination. Staff and experts consistently articulated a desire to enhance learning authenticity and integrity, while simultaneously expressing concern regarding sustainability and institutional support. Survey respondents comprised academic staff from the Faculty of Business and represented a mature cohort with 76% possessing over 10 years of higher education teaching experience. Despite this extensive experience, the data indicates that this does not automatically translate into confidence with GenAI-resilient AA design. Most respondents (84%) reported primary responsibility for module-level assessment design, while only 19% indicated regular involvement in programme-level assessment coordination. This pattern reflects a decentralised model where AA innovation is concentrated at the individual level rather than through coordinated programme leadership and design practices. Several survey respondents (55%) described feeling “a strong sense of individual responsibility and additional personal pressure” for maintaining standards without corresponding authority or support. The survey reveals a significant disparity between staff

familiarity with AA and the institutional support provided to implement it; 84% of respondents currently utilise AA, however, 49% of staff disagree that they have received any formal training or professional development in AA design (Figure 1). Survey responses revealed heavy reliance on informal learning networks, self-directed experimentation, and online resources.

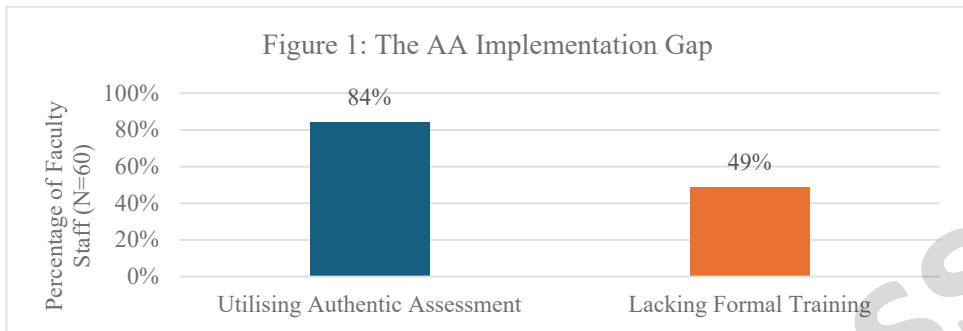
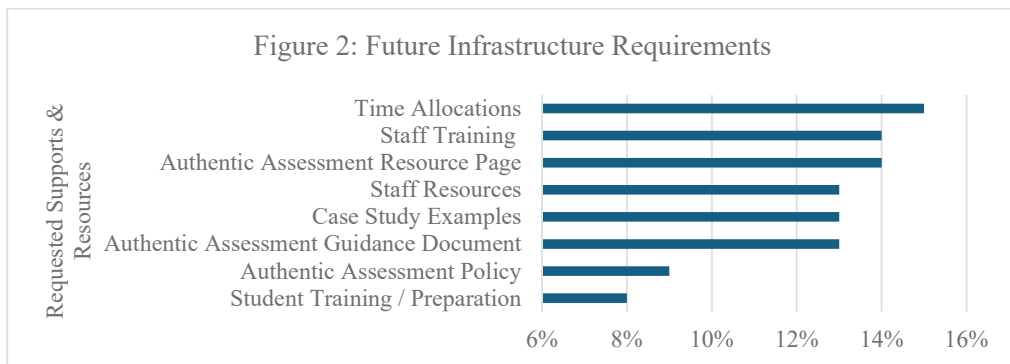


Figure 1: The Implementation Gap

Expert participants corroborated this gap, with one warning that without resourcing, authenticity risks becoming a rhetorical device rather than an operational reality: *"I'm cautious because... authenticity is a bit of a buzzword, and I think people do get buzzword fatigue, it becomes a panacea term, that we band here and say, look, AA is the magic bullet."* While staff strongly associate AA with enhanced engagement, its role as an integrity mechanism is viewed with caution. Respondents ranked "Reducing Academic Misconduct" only sixth among perceived benefits, suggesting that staff view integrity as a secondary by-product of good design rather than a primary driver of assessment reform. Expert participants emphasised the difficulty of verifying authorship when GenAI can simulate professional outputs, with one expert describing the challenge through the metaphor of "bread and jam": *"I've always called it the bread and jam syndrome. The student writes their own piece, and it can be tough like stale bread. And then you get this beautiful piece of jam in the middle of it. And you go, wow, that's interesting."* Across both datasets, there was profound scepticism toward detection technologies. Experts emphasised that compliant submissions may represent only a "semblance" of student learning rather than genuine cognitive engagement. Workload emerged as the most definitive, systemic barrier to sustainable reform. 61% of survey respondents explicitly identified "Time Allocations" as their primary challenge (Figure 2). Experts provided the "emotional weight" to this statistic, describing the demand for rapid, authentic feedback in large classes as a task that current teaching models simply cannot accommodate: *"18 hours of teaching, four or five modules. You're in a department with 50 students per module. [...] On the one hand, I'm talking about AA, and it's brilliant. And the next thing I talk about is feedback and getting feedback to your students within one day. [...] These are circles that we're not able to square off here."*

*Figure 2: Required Institutional Scaffolding*

Findings highlight widespread programme-level fragmentation. Uncoordinated module-level innovation generated student overload, duplication of assessment formats, and inconsistent verification practices. Students were frequently exposed to multiple concurrent industry projects, and reflective portfolios without curricular sequencing. Table 1 shows key patterns in AA implementation which indicate that challenges are systemic rather than individual.

Table 1. Empirical Patterns in AA Implementation

Pattern	Staff Accounts	Expert Accounts	Governance Implication
Implementation Gap	84% use AA; 49% no training	Unsustainable scaling	Capacity deficit
Labour Compression	Design overload	Structural misfit	Workload misalignment
Fragmented Capability	Informal learning	Weak scaffolding	Professional risk
Programme Disconnect	Assessment overload	“Fjord modules”	Coherence failure
Assurance Ambiguity	Limited detection trust	Judgement dependency	Integrity risk

The survey findings also reveal a pattern of fragmented excellence, where individual lecturers create high-impact AAs in isolation representing fjord modules (Lodge *et al.*, 2025). AA innovation is sustained through personal goodwill, informal mentoring, and moral obligation rather than through formal governance mechanisms. Without coordinated infrastructure, these exemplary practices fail to generate cumulative institutional benefit and remain vulnerable to staff burnout. Transitioning from fjord modules to a connected system requires moving beyond the myth of the “all-capable solo academic” toward an “assessment village” of shared responsibility Nicola-Richmond *et al.* (2025). The findings confirm that pedagogical commitment is necessary, but insufficient, and point to a shift toward programme-level

coordination and the need for institutional permissions to iterate and adapt in response to the GenAI condition. Without these permissions, AA innovation remains fragmented.

5. The Institutional Framework

This section consolidates the empirical findings with the conceptual foundations to advance an institutional theory of AA governance in GenAI-enabled environments. Rather than functioning as a pedagogical toolkit, the framework proposed by the authors, is positioned as a form of governance infrastructure through which professional judgement, organisational learning, and quality assurance are systematically coordinated. The nested architecture illustrated in the framework (Figure 3) depicts how pedagogical design, professional capability, governance alignment, and evidence-informed review operate as interdependent components of a self-reinforcing system. Unlike linear implementation models, the flywheel conceptualises reform as an iterative, cumulative process. Momentum is generated through reciprocal reinforcement rather than through periodic intervention. The proposed framework is operationalised through dedicated institutional infrastructures, specifically Digital Learning Hubs and Assessment Sandboxes. These are not merely conceptual spaces but the operational core of the framework, providing structured environments for collaborative design, resource sharing, and safe experimentation without regulatory risk. By embedding innovation within these supported systems, institutions can reduce reliance on “heroic” individual goodwill and ensure that reform is both scalable and sustainable (Gunder *et al.*, 2025; Nicola-Richmond *et al.*, 2025).



Figure 3: Institutional Framework for Authentic Assessment Governance

6. Implications and Conclusion

The findings and framework developed in this study have significant implications for institutional leadership, quality assurance policy, professional practice in GenAI-enabled higher education and AA practices. Assessment reform must be embedded within institutional strategies and performance frameworks and must include explicit recognition of AA design and evaluation efforts within workload allocation models and academic progression pathways. In order to reduce programme-wide fragmentation and duplication, institutions must establish formal mechanisms for cross-programme coordination such as ongoing programme-level assessment leadership roles, supported by time allocation and professional authority. Investment in digital learning hubs and assessment sandboxes as core academic infrastructure must be prioritised. Regulatory frameworks must evolve to recognise adaptive judgement as a legitimate component of quality assurance. This involves creating space within accreditation and review processes for iterative AA redesign, contextual variation, and negotiated standards. National agencies and professional bodies should support sector-wide learning by facilitating the exchange of infrastructural models, such as the institutional framework presented in this study, as well as evaluation tools, and governance practices. Professional development must be integrated into workload models to achieve sustained impact. Institutions should develop accredited pathways in assessment governance and GenAI-resilient AA design. Re-design clinics, peer observation of assessment practices and programme-based communities of practice should be prioritised. Academics should be actively encouraged and supported through these initiatives to engage in AA design that incorporates staged, dialogic, and reflective authentic tasks. By adopting the proposed framework, universities can transform technological disruption into an opportunity for institutional learning, enhanced validity, and professional resilience.

References

- Corbin, T., Bearman, M., Boud, D., & Dawson, P. (2025). The wicked problem of AI and assessment. *Assessment & Evaluation in Higher Education*, 1–17. <https://doi.org/10.1080/02602938.2025.2553340>
- Gunder, A., Vito, M., Ramirez, M., Ford, C. (2025). Beyond AI Awareness: How Institutions Are Strategically and Pedagogically Embedding AI Skills Development. *Ubiquity Proceedings*, 6(1): 14. DOI: <https://doi.org/10.5334/uproc.182>
- Lehane, S., Wright, A., & Fenton, P. (2024). Improving academic integrity through AA design. *Irish Educational Studies*, 1–23. <https://doi.org/10.1080/03323315.2024.2441164>
- Lehane, S., Wright, A., & Fenton, P. (2025). Utilising student reflections to inform authentic assessment design. *European Journal of Higher Education*, 1–27. <https://doi.org/10.1080/21568235.2025.2561157>
- Lodge, J. M., Bearman, M., Dawson, P., Gniel, H., Harper, R., Liu, D., McLean, J., Ucnik, L., & Associates. (2025). Enacting assessment reform in a time of artificial intelligence. Tertiary Education Quality and Standards Agency, Australian Government.

Nicola-Richmond, K., Dawson, P., Helen, P., & Macfarlane, S. (2025). *It takes a village. Program-wide approaches to redesigning assessment in a time of generative artificial intelligence (GenAI). (Version 1).* Deakin University.
<https://hdl.handle.net/10779/DRO/DU:30341809.v1>

Article in Press