



AI-Aligned Assessment Practices for a Future-Ready Higher Education System

Simonne Stellenboom* and Buhleni Ncube

Stadio Higher Education, South Africa

Abstract

The rapid integration of generative artificial intelligence (GenAI) into higher education has prompted institutions to reconsider traditional assessment methods. As students increasingly use GenAI for brainstorming, drafting, and problem-solving, assessments based on memorisation or predictable outcomes no longer reliably measure independent learning or authentic skill development. Experimental evidence confirms that AI-generated submissions can achieve pass-level grades and that markers cannot reliably distinguish AI-authored work from student work, even in tasks designed to be authentic. This study conducts a systematic review of 25 peer-reviewed articles (2022–2026) examining generative AI, assessment, and academic integrity in higher education, complemented by practice-based illustrations from Commerce modules at a South African private higher education institution. Addressing the shortage of integrated, task-level guidance for assessment redesign, the review proposes the AI-Aligned Assessment Framework: five interconnected principles—Critical Judgement, Ethical AI Use, Scenario-Based Application, Learning Process Visibility, and Transparent Student–AI Collaboration. The framework treats AI capability as a moving target rather than a fixed boundary and is designed to be scalable across disciplines and institutional contexts, including those characterised by high student-to-staff ratios. Rather than designing primarily to detect or exclude AI use, the framework shifts assessment design towards generating credible evidence of student reasoning, accountability, and contextual application. The paper contributes a practical, discipline-adaptable framework that supports academic integrity, authentic learning, and graduate preparedness in AI-enhanced higher education.

Keywords: academic integrity, authentic assessment, AI literacy, graduate preparedness, higher education

1. Introduction

Generative artificial intelligence has moved rapidly from a novel technology to an embedded feature of higher education. Students use GenAI tools to support writing, brainstorming, and problem-solving across disciplines (Fern, 2024; Evangelista, 2025), which complicates the relationship between submitted work and evidence of individual learning (Evangelista, 2025).

Many traditional assessment formats - closed-book recall tests, standard essays, generic case questions - assume that students already possess the knowledge needed to produce a competent answer. That assumption is increasingly untenable: experimental evidence shows GenAI can complete traditional essay assessments competently and quickly and can also satisfy the requirements of tasks commonly regarded as authentic (Klyshbekova & Abbott, 2024; Kofinas et al., 2025).

GenAI capability should not be treated as uniform or static. Performance varies across systems, model generations, and task types, and this pattern has been documented as a “jagged technological frontier”: AI assistance improves performance on some tasks while degrading it on others that appear similarly difficult (Dell’Acqua et al., 2026). Earlier studies identified specific limitations - for instance, Klyshbekova and Abbott (2024) found ChatGPT produced generic, inadequately supported essay content with fabricated references - but such findings describe a particular model at a particular point in time rather than a stable boundary. Assessment design should therefore treat AI capability as continuously evolving rather than rely on today’s limitations as a defence.

The evidence that conventional assessment is vulnerable is concrete. In experiments across two UK universities, markers could not reliably distinguish student-authored work from GenAI-modified or GenAI-generated submissions and increasing the apparent authenticity of a task had little effect on detection. Kofinas et al. (2025) conclude that authentic assessment alone cannot safeguard integrity and recommend greater attention to assessment design, learning process, and workplace-relevant performance. AI-detection systems offer no dependable alternative: they generate false positives and negatives, encourage suspicion-based marking, and create an adversarial assessment climate rather than improving validity (Perkins et al., 2025).

These challenges carry weight in African and other Global South contexts. Maluleke’s (2025) systematic review of African higher education identifies four recurring challenges: academic-integrity risk, data-privacy concern, digital and infrastructural inequality, and inadequate institutional readiness. African institutions therefore need context-specific approaches rather than imported models built for different environments. The policy response remains uneven globally: Moorhouse et al. (2023) found fewer than half of the world’s 50 highest-ranked universities had public GenAI assessment guidance, typically limited to integrity, design, and communication. Nicola-Richmond et al. (2025) add that redesign is resource-intensive - academics and students recognise the need for change but cite complexity and expertise as barriers, concluding that meaningful redesign “takes a village” and favours programme-wide, collaborative approaches over isolated effort.

Despite rapid growth in this literature, it remains fragmented, concentrated on misconduct, detection, or individual strategies (oral exams, authentic tasks, disclosure) without an integrated system connecting them (Gao, 2024). Early institutional responses called for proactive combinations of policy, staff training, and assessment adaptation rather than reliance on prevention alone (Cotton et al., 2023), and a broader systematic review of 41 studies confirms that GenAI simultaneously enhances engagement and intensifies integrity risk, reinforcing the need for an integrated design response rather than piecemeal guidance (Bittle & El-Gayar, 2025). Perkins (2023) adds a foundational distinction that runs through this framework: AI use itself is not inherently misconduct, but its concealment is - a position that shifts institutional attention from banning tools to requiring disclosure, and one that Chan’s (2023) institution-wide policy framework extends by linking pedagogical, governance, and operational responses rather than leaving assessment-level guidance to individual educators.

This paper makes four contributions: it demonstrates why traditional formats can no longer be assumed to provide reliable evidence of learning; it identifies the shortage of integrated, task-level guidance for educators; it argues that sustainable assessment requires shifting from detecting or excluding AI towards designs in which students remain intellectually responsible and accountable; and it synthesises current evidence into the AI-Aligned Assessment Framework, comprising Critical Judgement, Ethical AI Use, Scenario-Based Application, Learning Process Visibility, and Transparent Student-AI Collaboration. The framework offers a design-oriented structure adaptable across disciplines and institutional contexts rather than a single prescribed format or fixed permitted level of AI use.

1.1 Research Question

This study is guided by one question: How can higher education assessment practices be reimaged to remain valid, ethical, and meaningful in an AI-enhanced learning environment? Validity requires assessment to measure students' own knowledge, skills, and judgement rather than what AI can generate on their behalf. Ethics requires assessment design that is fair and transparent, supporting rather than criminalising responsible AI engagement. Meaningfulness requires preparing students for complex, digitally mediated professional environments by developing transferable skills that remain valuable alongside AI.

2. Methodology

This study employed a systematic literature review, following PRISMA guidelines (Page et al., 2021), to synthesise evidence on generative AI's implications for higher education assessment and to develop an evidence-informed conceptual framework. A systematic approach was selected over a narrative review because the literature has expanded rapidly and a structured, transparent process reduces selection bias and improves reproducibility. Figure 1 documents the review process.

Searches were conducted across Scopus, Google Scholar, and discipline-specific databases using terms combining generative AI, large language models, and ChatGPT with assessment, academic integrity, and higher education, supplemented by backward and forward citation searching. The primary search string was: ("generative AI" OR "ChatGPT" OR "large language model") AND ("assessment" OR "academic integrity") AND ("higher education" OR "university"). Studies were included if peer-reviewed, published in English between 2022 and 2026, focused on higher education, and examined GenAI in relation to assessment, integrity, or assessment design; studies on primary/secondary education, unrelated AI applications, or non-peer-reviewed formats were excluded. The first author conducted the primary database search.

The second author conducted a supplementary search focused specifically on African and South African higher education contexts, broadening the geographic representation of the evidence base. Searches were conducted between January and March 2026. Both authors independently screened titles and abstracts against the inclusion and exclusion criteria; they resolved disagreements through discussion until consensus was reached. Both authors conducted the full-text review and jointly agreed on final inclusion decisions. A formal risk-of-bias and quality appraisal was applied to every full-text article against four criteria: clarity of research question or purpose, appropriateness of methodology, rigour of evidence or argumentation, and relevance to the five thematic areas. Studies were retained where they met at least three of the four criteria; no study was excluded solely on methodological grounds where its conceptual contribution remained substantive. The outcomes of this

appraisal are reflected in the exclusion counts reported in Figure 1, where 12 records were removed for not being peer-reviewed and a further 13 for insufficient assessment focus or unmet quality criteria.

Following title/abstract screening and full-text review, 25 studies were retained for thematic synthesis. Initial codes covering AI capability, assessment validity, integrity, and implementation challenges were iteratively refined into five recurring themes - Critical Judgement, Ethical AI Use, Scenario-Based Application, Learning Process Visibility, and Transparent Student–AI Collaboration - which were then synthesised into the AI-Aligned Assessment Framework.

To strengthen practical relevance, the synthesis was complemented by practice-based illustrations drawn from undergraduate Commerce modules at a South African private higher education institution. These illustrations are worked examples, not empirical case-study findings; the authors developed them through professional practice and refined them through lecturer feedback and moderation across assessment cycles. The institution is not named in keeping with standard practice for practice-informed research of this kind. Because no student data were collected or analysed, institutional ethical clearance was not required.

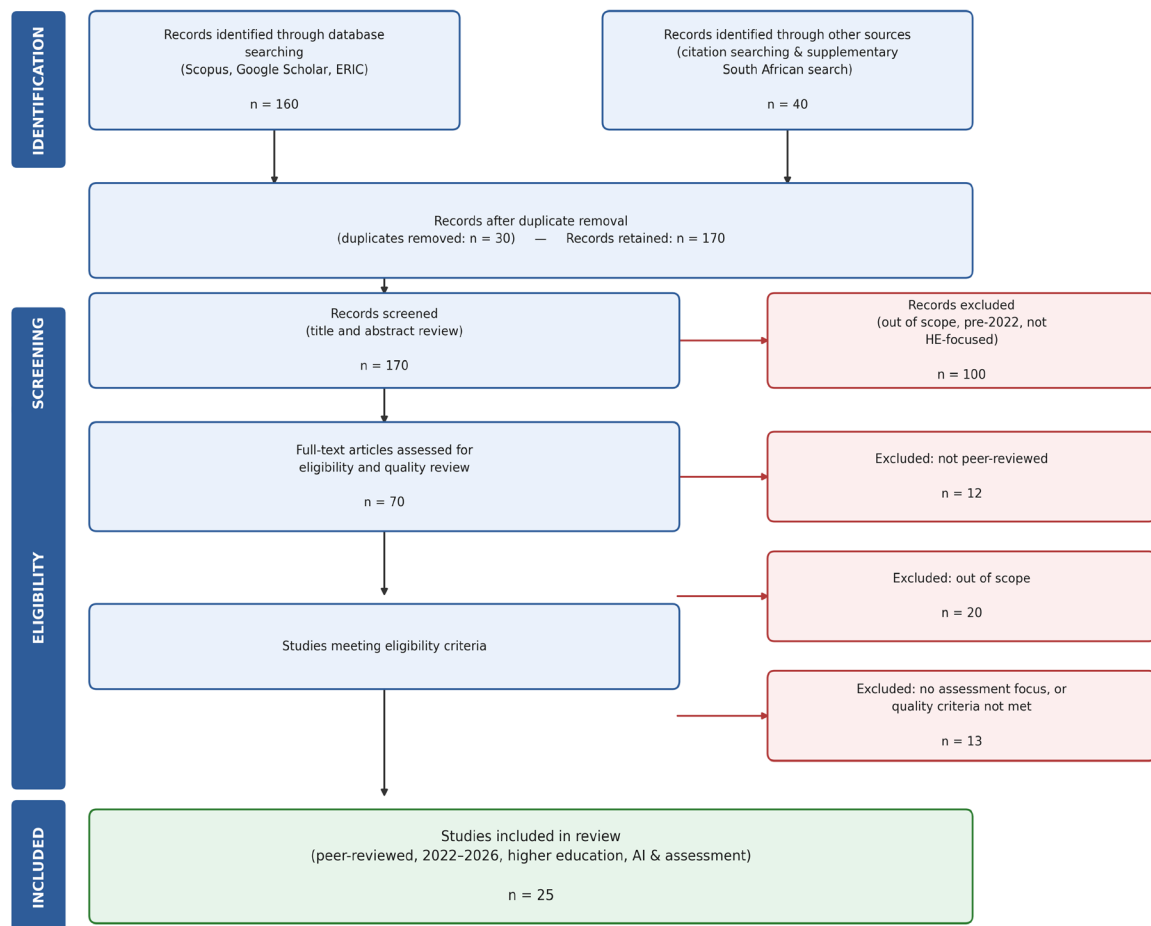


Figure 1. PRISMA-ScR Flow Diagram of Study Selection Process

Source: Authors, adapted from PRISMA 2020 (Page et al., 2021)

2.1 Limitations

This review has several limitations. The restriction to English-language publications between 2022 and 2026 may introduce publication bias, potentially excluding relevant work published in other languages or prior to 2022. The rapid pace of GenAI development means some findings—particularly those describing specific AI capabilities—may already be partially superseded. The practice-based illustrations are drawn from a single institutional context in South Africa and may not transfer without adaptation to other disciplines or institutional settings. Future research should empirically validate the framework across disciplines, institutions, and linguistic contexts.

Table 1 presents the characteristics of the 25 studies included in the review, mapped against the five framework principles. The table reflects the geographic distribution of the evidence base, the range of methodological approaches represented, and the thematic alignment of each study with the principles developed through the synthesis.

Table 1: Characteristics of Included Studies Mapped to Framework Principles

Author(s) & Year	Country	Methodology	Principles(s)
Bittle & El-Gayar (2025)	USA	Systematic review	P1–P5
Chaka et al. (2024)	South Africa	Content analysis	P2, P5
Chan (2023)	Hong Kong	Mixed methods	P2
Cotton et al. (2023)	UK	Experimental	P2, P5
Dell'Acqua et al. (2026)	USA	Field experiment	P1, P3
Eaton (2023)	Canada	Conceptual	P2, P5
Evangelista (2025)	Philippines	Literature review	P2, P3
Fern (2024)	USA	Conceptual	P4, P5
Finkel-Gates (2025)	USA	Literature review	P4, P5
Francis et al. (2025)	UK	Literature review	P2, P4
Gao (2024)	China	Literature review	P1, P3
Gonsalves (2025)	Canada	Conceptual	P3
Gundu (2025)	South Africa	Systematic review	P1, P3
Klyshbekova & Abbott (2024)	Kazakhstan/UK	Experimental	P1
Kofinas et al. (2025)	UK	Experimental	P1, P4
Lehane & Wright (2024)	Ireland	Case study	P3, P4
Lehane et al. (2025)	Ireland	Thematic analysis	P3, P4
Maluleke (2025)	South Africa	Systematic review	P2, P4
Mithi et al. (2024)	South Africa	Qualitative	P4
Moorhouse et al. (2023)	Hong Kong	Document analysis	P2
Ncube et al. (2026)	South Africa	Systematic review	P4, P5
Nicola-Richmond et al. (2025)	Australia	Literature review	P4
Page et al. (2024)	USA	Conceptual framework	P1

Author(s) & Year	Country	Methodology	Principles(s)
Perkins (2023)	Australia	Conceptual	P2, P5
Perkins et al. (2024)	Australia	Conceptual framework	P2
Perkins et al. (2025)	Australia	Conceptual framework	P2, P5
Salinas-Navarro et al. (2024)	Mexico/UK	Ethnographic	P3, P5
Wang & Sun (2025)	China	Systematic review	P3, P5

Note. P1 = Critical Judgement; P2 = Ethical AI Use; P3 = Scenario-Based Application; P4 = Learning Process Visibility; P5 = Transparent Student–AI Collaboration.

Source: Authors

3. Literature Review

3.1 The Expanding Capability of Generative AI

Early experimental work established that GenAI could complete assessment tasks competently. Klyshbekova and Abbott (2024) found that ChatGPT produced an academic essay quickly, but the work was generic and descriptive, used fabricated references, and scored substantially higher when graded by the system itself than by the researchers. Kofinas et al. (2025) extended this using student-authored, AI-modified, and AI-generated submissions across two UK business schools: markers could not reliably distinguish between categories, and increasing task authenticity had little effect on detection—authenticity alone therefore cannot be assumed to protect integrity.

The relationship between AI capability and task complexity is not linear. Dell’Acqua et al. (2026) introduce the “jagged technological frontier”: in a preregistered field experiment with 758 knowledge workers, GPT-4 substantially improved productivity and quality on tasks within its capability frontier but reduced correctness on a task situated just beyond it, despite similar apparent difficulty. Because this boundary is difficult to predict in advance and continues to shift, assessment design cannot rely on identifying tasks that are permanently “AI-resistant.” Bloom’s Taxonomy remains useful for aligning cognitive demand with learning outcomes but should not be treated as a stable line separating human and AI capability (Page et al., 2024).

3.2 From AI Limitations to Contextual and Experiential Assessment

Designing around what AI currently cannot do is consequently unstable. The more durable response, evident across the literature, is to increase the contextual and experiential density of assessment rather than search for AI-proof cognitive levels. Gonsalves (2025) argues that authenticity alone is insufficient and proposes five dimensions of context—practical, situational, experiential, interdisciplinary, and ethical—that require students to identify, interpret, and integrate relevant information rather than merely produce a plausible answer. Contextualisation is not achieved by adding a company name to a generic question; its value lies in requiring genuine interpretation of circumstance.

Salinas-Navarro et al. (2024) similarly reposition GenAI from an “agent-to-write” or “agent-to-answer” into a support for human-centred experiential learning, describing a continuous cycle of concrete experience, reflection, conceptualisation, and active experimentation within constructively aligned assessment. Under this model, GenAI can support learning activities without displacing the accountability that connects intended outcomes, teaching activities,

and assessment evidence. Security therefore does not arise from a task appearing authentic or difficult; it arises from requiring students to exercise judgement within a specific context, explain their reasoning, and remain accountable for their conclusions (Gonsalves, 2025).

3.3 Designing Around Human Learning Rather Than AI Resistance

Three enduring properties emerge from this literature. First, assessment should be grounded in sufficiently specific and evolving context—current conditions, local regulation, or a student’s own prior decisions—requiring students to determine what information is relevant rather than assuming AI cannot access it. Second, assessment should require accountable judgement: comparing alternatives, identifying trade-offs, and justifying professional or ethical consequences. Third, assessment should generate visible evidence of the learning process—multi-stage work, oral defence, reflection, and continuous feedback—since process evidence is harder to substitute wholesale than a finished product (Ncube et al., 2026). Process visibility is not surveillance; its purpose is educational evidence of how a problem was defined, evidence selected, feedback used, and AI suggestions accepted or rejected. Taken together, the central assessment question shifts from “Can AI complete this task?” to “What evidence does this assessment provide that the student engaged in the required thinking?”—the theoretical foundation for the framework below.

4. The AI-Aligned Assessment Framework

The thematic synthesis identified five interconnected principles that, together, shift assessment from detecting AI use towards designing learning experiences that promote judgement, accountability, and meaningful engagement with AI (Figure 2). The framework does not assume any task is permanently AI-resistant; it is organised around enduring educational purposes rather than the present limitations of any model class.

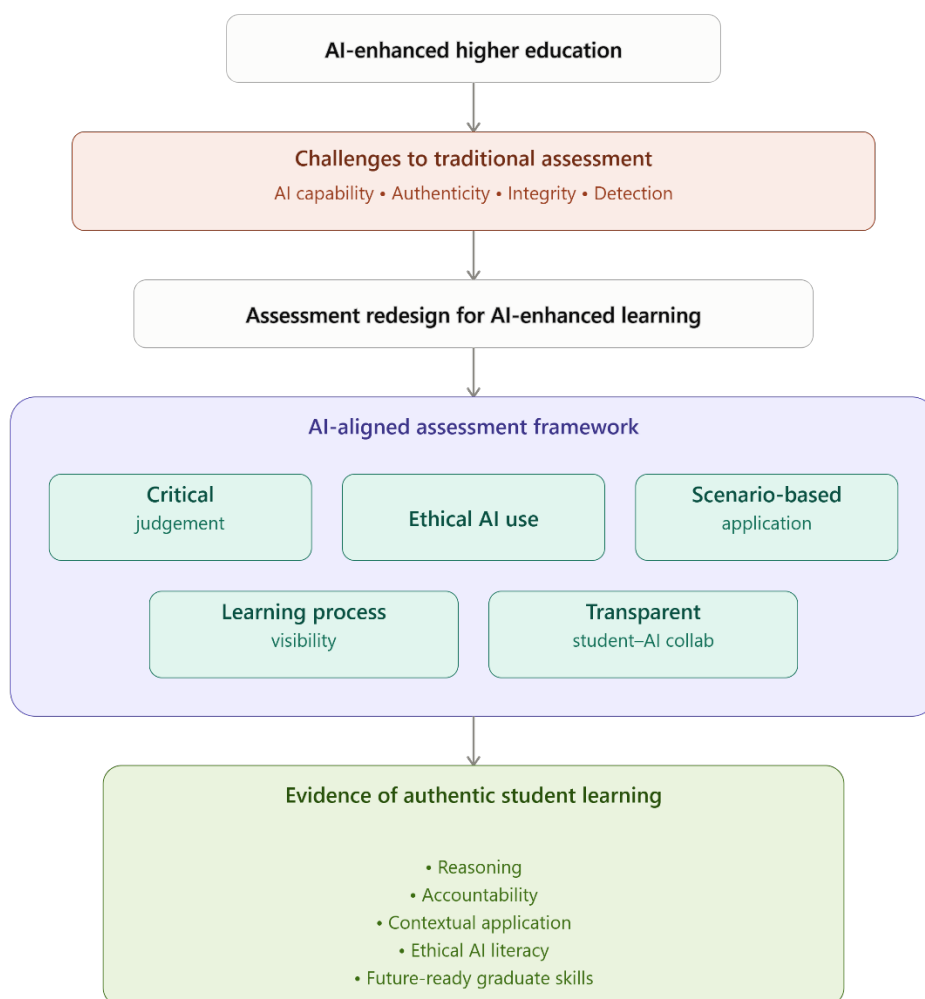


Figure 2. *The AI-Aligned Assessment Framework*
Source: Authors

The framework comprises five complementary principles: (1) Critical Judgement, (2) Ethical AI Use, (3) Scenario-Based Application, (4) Learning Process Visibility, and (5) Transparent Student–AI Collaboration. Collectively, these principles promote valid assessment by ensuring that students demonstrate reasoning, accountability, contextual application, and responsible engagement with AI throughout the learning process.

4.1 Principle One: Critical Judgement

AI-aligned assessment should require students to interpret information, compare alternatives, and justify decisions—not because AI cannot perform higher-order tasks, but so students remain accountable for the judgement in their submitted work. Because task difficulty does not predict AI performance in advance (Dell’Acqua et al., 2026), tasks should require students to justify decisions using disciplinary evidence, compare alternative solutions, acknowledge uncertainty, identify where AI output may be unreliable, and defend a professional recommendation, with marks rewarding the quality of reasoning rather than a polished final answer.

In a Commerce context, this principle replaces a question such as “What is the formula for net present value?” with a task requiring students to evaluate competing investment proposals under financial, ethical, and strategic constraints, calculate and interpret the relevant measures, and justify a recommendation. AI may support calculations or generate initial

alternatives, but the student remains responsible for verifying information and defending the final decision.

4.2 Principle Two: Ethical AI Use

Prohibition is not a sustainable response to widespread GenAI use. The AI Assessment Scale (AIAS), introduced by Perkins et al. (2024) and subsequently refined, provides a structured, non-hierarchical approach ranging from no AI use to full AI exploration, aligning permitted use with learning outcomes and attending to equity, privacy, and disclosure (Perkins et al., 2025). Eaton (2023) frames the broader shift as a post plagiarism era, in which hybrid human–AI work becomes a normal condition of academic life rather than an exception to be policed, requiring ethical frameworks built on transparency rather than detection. Operationalising this principle requires assessment-specific guidance on permitted and prohibited uses, student AI-use declarations, responsibility for verifying AI-generated content, and clearly communicated consequences for concealed use. Salinas-Navarro et al. (2024) add that ethical use should prioritise human control and higher-order thinking rather than treating AI output as a finished product.

General institutional policy is insufficient where students cannot determine how it applies to a specific task. South African evidence shows uneven institutional responses to GenAI, reinforcing the need for assessment-level guidance that is transparent and locally appropriate (Chaka et al., 2024). Ethical AI Use therefore extends beyond preventing misconduct: it prepares students to evaluate AI critically and take responsibility for AI-supported work in professional contexts, consistent with calls for GenAI to complement rather than replace human expertise as institutions balance innovation against integrity (Francis et al., 2025).

4.3 Principle Three: Scenario-Based Application

Scenario-Based Application requires students to apply disciplinary knowledge within situated problems involving incomplete information and competing constraints. Its value does not lie in assuming complex scenarios cannot be completed by AI, but in creating opportunities to evaluate how students interpret context, manage uncertainty, and justify a position. Effective scenarios include simulated workplace tasks, cases with incomplete or contradictory evidence, decisions involving ethical trade-offs, multiple defensible solution paths, and a requirement to critique an AI-generated recommendation (Salinas-Navarro et al., 2024); adaptive strategies such as randomised questioning, project-based tasks, and open-book formats offer practical starting points, particularly in online and distance-learning contexts (Gundu, 2025). Empirical evidence from a cohort of 128 final-year business students supports the approach: deliberately designed authentic assessment strengthened engagement, employability, and adherence to integrity expectations, provided implementation responded to student expectations through clear guidance and staff development rather than simply labelling tasks authentic (Lehane & Wright, 2024; Lehane et al., 2025).

For example, a Commerce assessment might present a client with R500,000 to invest across options with different risk profiles and ethical implications; students analyse the evidence, account for the client's constraints, and remain accountable for their recommendation, including evaluating any AI-supported analysis they used. The same logic adapts across disciplines: Humanities students might compare competing historical interpretations and account for source bias; Science students might evaluate contradictory results and propose a justified course of action; Health Science students might recommend an ethically defensible response to a complex scenario while acknowledging the limits of available evidence.

4.4 Principle Four: Learning Process Visibility

This principle shifts attention from the final product to evidence of how understanding and judgement developed. Because markers cannot reliably distinguish AI-generated from student-authored work, and authenticity alone does not improve detection (Kofinas et al., 2025), assessment should make the learning process itself visible—through staged submissions, annotated drafts, documented AI use and evaluation, oral explanations of key decisions, peer feedback, and concise process portfolios. The purpose is not continuous surveillance, but evidence of how a problem was defined, evidence selected, and AI suggestions accepted or rejected.

The workload implications must be acknowledged. Nicola-Richmond et al. (2025) found that academics and students recognise the need for change but cite complexity and resourcing as pervasive concerns, recommending collaborative, programme-wide implementation over isolated redesign. South African evidence reinforces this from the formative side: GenAI can strengthen feedback and responsiveness in formative assessment, but only where educator capacity and clear ethical guidance accompany implementation (Mithi et al., 2024). Practical strategies include sharing design expertise across programme teams, coordinating staged assessment across modules, using low- or zero-weighted developmental activities, and introducing the principle incrementally. Not every stage requires formal grading; drafts and reflections can carry low or no marks while still generating useful evidence of engagement.

4.5 Principle Five: Transparent Student-AI Collaboration

This principle positions AI as a learning support used openly, critically, and accountably. Students should be able to explain what AI contributed, how outputs were verified, and where their own judgement shaped the final work—consistent with Salinas-Navarro et al.'s (2024) repositioning of GenAI as a support for experiential learning rather than a substitute for it, and with qualitative evidence that process-oriented assessment paired with explicit disclosure helps ensure AI supplements rather than substitutes student learning (Finkel-Gates, 2025). The refined AIAS recommends transparency regarding AI assistance and advises teaching disclosure, privacy, and intellectual property explicitly, functioning as a communication and redesign tool rather than a restriction mechanism (Perkins et al., 2025). Ncube et al. (2026) similarly identify disclosure policies and human–AI collaboration as central to responsible assessment, and a broader review of GenAI in digital education cautions that these benefits depend on sustained human oversight, data protection, and ethical governance rather than assuming collaboration is safe by default (Wang & Sun, 2025).

Transparent collaboration may require students to document the tool used, its purpose, how outputs were checked, and which aspects of the submission reflect their own judgement. The required documentation should be proportionate: a short AI-use statement may suffice for routine tasks, while a significant project may warrant an annotated output or reflection. Students remain responsible for all submitted material, including inaccurate or fabricated AI content; transparency does not transfer accountability to the technology.

4.6 The Framework as an Integrated Design Model

Although presented separately, the five principles operate as an integrated model rather than a checklist. Critical Judgement is strengthened by context-rich scenarios; Ethical AI Use establishes the boundaries within which AI may be used; Process Visibility provides evidence of how judgement developed; Transparent Collaboration shows how AI supported, but did not replace, the student's contribution. The framework represents a shift in assessment

philosophy: rather than designing primarily to identify AI use, educators design assessment to ensure students remain intellectually responsible for learning regardless of AI involvement. Table 2 summarises the framework.

Table 2: Summary of the AI-Aligned Assessment Framework

Framework principle	Primary purpose	Guiding question for educators	Evidence of student learning
Critical Judgement	Disciplinary reasoning and decision-making	Can students justify and defend their conclusions?	Evaluation, reasoning and defensible judgement
Ethical AI Use	Responsible and accountable AI engagement	Is permitted AI use clearly defined and ethically applied?	Accountability, verification and ethical decision-making
Scenario-Based Application	Application within context	Can students apply knowledge under uncertainty and competing constraints?	Contextual interpretation and professional judgement
Learning Process Visibility	Evidence of learning development	Can the development of the student's thinking be observed?	Reflection, iteration, feedback response and decision development
Transparent Student-AI Collaboration	Critical AI literacy	Can students explain, evaluate and take responsibility for their AI use?	Disclosure, verification, revision and responsible collaboration

Source: Authors

Figure 3 traces how the initial coding categories from the thematic synthesis converged into five recurring themes, which were then consolidated into the AI-Aligned Assessment Framework.

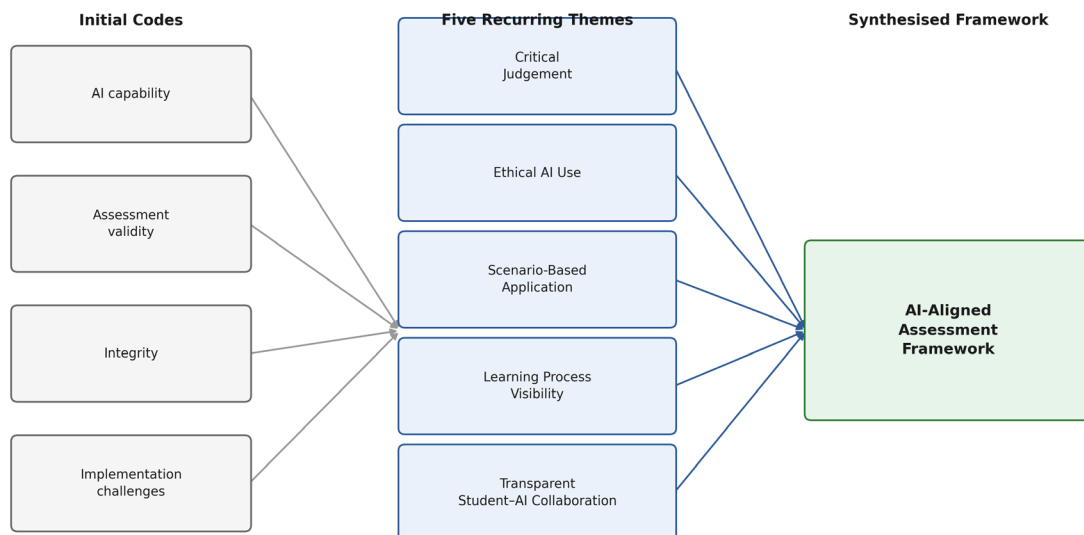


Figure 3. From Literature Themes to the AI-Aligned Assessment Framework

Source: Authors

5. Discussion

The AI-Aligned Assessment Framework shifts assessment from preventing or detecting AI use towards designing assessments that provide credible evidence of learning regardless of whether AI is available. This distinction matters because AI performance does not follow a predictable cognitive boundary: capability is uneven across tasks, and two tasks that appear equally difficult to educators may produce very different human–AI performance outcomes (Dell’Acqua et al., 2026). The framework therefore complements rather than replaces the AIAS: where the AIAS specifies acceptable levels of AI use, the AI-Aligned Assessment Framework specifies the characteristics—judgement, accountability, contextual application, visible process, and transparent collaboration—that should be embedded within the task itself.

The framework also extends authentic-assessment and experiential-learning literature (Salinas-Navarro et al., 2024) but cautions against treating authentic or experiential tasks as automatically secure. Their value lies in requiring meaningful engagement, reflection, and evaluative judgement, which is why Scenario-Based Application must be combined with Critical Judgement and Process Visibility rather than relied upon alone.

The five principles are discipline-agnostic but not discipline-neutral: their implementation must reflect the knowledge practices of each field. Critical Judgement in Finance may involve evaluating investment alternatives; in Health Sciences, weighing evidence and patient circumstances; in Engineering, balancing safety, cost, and sustainability; in the Humanities, interpreting contested evidence. The framework provides a common design logic while allowing evidence of learning to differ by discipline.

Implementation must also acknowledge institutional capacity. Given that redesign is resource-intensive (Nicola-Richmond et al., 2025), sustainable implementation may involve incremental integration into existing assessments, selective staging for significant tasks, shared scenarios and rubrics across teaching teams, low-stakes developmental stages, and programme-level coordination of AI-use guidance. The framework’s central contribution is to reframe assessment from a technology-centred problem into a learning-centred design challenge: the relevant question is not whether AI was used, but whether the assessment provides convincing evidence that the student engaged in the required thinking and can explain AI’s role in producing the work.

6. Conclusion

The widespread adoption of generative AI has altered the assumptions underpinning higher education assessment. As AI systems become increasingly capable of producing academically credible work, approaches that depend primarily on preventing or detecting AI use are becoming less sustainable, and the jagged nature of AI capability complicates any attempt to define permanently AI-resistant tasks. Assessment design must therefore move beyond assumptions about current technological boundaries and focus on the evidence students provide of their own learning and judgement.

Drawing on a systematic review of the literature, this paper proposes the AI-Aligned Assessment Framework, comprising five interconnected principles—Critical Judgement, Ethical AI Use, Scenario-Based Application, Learning Process Visibility, and Transparent Student–AI Collaboration—that together place student reasoning, contextual application, accountability, and AI literacy at the centre of assessment. The framework recognises that

redesign requires expertise, collaboration, and institutional support; programme-wide coordination and incremental implementation offer a more sustainable route than expecting individual educators to redesign assessment in isolation. It is intended to guide rather than prescribe, adaptable across disciplines, qualification levels, and institutional contexts.

Future research should empirically examine the framework across disciplines and institutions, evaluating its effects on assessment validity, student learning, educator workload, AI literacy, and equity of access, and refining it as AI capabilities and institutional policy continue to develop. The aim is not to compete with artificial intelligence or exclude it from learning, but to design assessments that keep students intellectually responsible, ethically accountable, and professionally prepared for an AI-enhanced world.

Acknowledgment

This paper was virtually presented at the 11th International Conference on Advanced Research in Teaching and Education (ICATE 2026), Kuala Lumpur, Malaysia, 26–28 June 2026. The authors thank STADIO Higher Education for its institutional support.

References

- Bittle, K., & El-Gayar, O. (2025). Generative AI and academic integrity in higher education: A systematic review and research agenda. *Information*, 16(4), 296. <https://doi.org/10.3390/info16040296>
- Chaka, C., Shange, T., Nkhobo, T., & Hlatshwayo, V. (2024). An environmental review of the generative artificial intelligence policies and guidelines of South African higher education institutions: A content analysis. *International Journal of Learning, Teaching and Educational Research*, 23(12), 487–511. <https://doi.org/10.26803/ijlter.23.12.24>
- 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, 38. <https://doi.org/10.1186/s41239-023-00408-3>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). 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>
- Dell'Acqua, F., McFowland III, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K. C., Rajendran, S., Krayner, L., Candelon, F., & Lakhani, K. R. (2026). Navigating the jagged technological frontier: Field experimental evidence of the effects of artificial intelligence on knowledge worker productivity and quality. *Organization Science*, 37(2), 403–423. <https://doi.org/10.1287/orsc.2025.21838>
- Eaton, S. E. (2023). Postplagiarism: Transdisciplinary ethics and integrity in the age of artificial intelligence and neurotechnology. *International Journal for Educational Integrity*, 19, 23. <https://doi.org/10.1007/s40979-023-00144-1>
- 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(1), Article ep559. <https://doi.org/10.30935/cedtech/15775>
- Fern, J. (2024). A more-than-human ecology: Evolving generative artificial intelligence in higher education. *Education Sciences*, 14(10), Article 1102. <https://doi.org/10.3390/educsci14101102>

- Finkel-Gates, A. (2025). ChatGPT in academic assessments: Upholding integrity. *Journal of Learning Development in Higher Education*, (36). <https://doi.org/10.47408/jldhe.vi36.1491>
- Francis, N. J., Jones, S., & Smith, D. P. (2025). Generative AI in higher education: Balancing innovation and integrity. *British Journal of Biomedical Science*, 81, 14048. <https://doi.org/10.3389/bjbs.2024.14048>
- Gao, L. (2024). Impact of ChatGPT on academic integrity and assessment effectiveness for e-learning in higher education: The need for redesigning assessment practices. *Communications in Humanities Research*, 45, 40–44. <https://doi.org/10.54254/2753-7064/45/20241697>
- Gonsalves, C. (2025). Contextual assessment design in the age of generative AI. *Journal of Learning Development in Higher Education*, (34). <https://doi.org/10.47408/jldhe.vi34.1464>
- Gundu, T. (2025). Strategies for e-assessments in the era of generative artificial intelligence. *The Electronic Journal of e-Learning*, 22(7), 40–50. <https://doi.org/10.34190/ejel.22.7.3477>
- Klyshbekova, M., & Abbott, P. (2024). ChatGPT and assessment in higher education: A magic wand or a disruptor? *The Electronic Journal of e-Learning*, 21(5), 407–419. <https://doi.org/10.34190/ejel.21.5.3114>
- Kofinas, A., Tsay, C. H., & Pike, D. A. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology*, 56(6), 2522–2549. <https://doi.org/10.1111/bjet.13585>
- Lehane, S., & Wright, A. (2024, June 18–21). Designing authentic assessment to improve academic integrity [Paper presentation]. *10th International Conference on Higher Education Advances (HEAd'24)*, Valencia, Spain. <https://doi.org/10.4995/HEAd24.2024.17136>
- Lehane, S., Wright, A., & Fenton, P. (2025, June 17–20). Student perspectives on improving academic integrity through authentic assessment design [Paper presentation]. *11th International Conference on Higher Education Advances (HEAd'25)*, Valencia, Spain. <https://doi.org/10.4995/HEAd25.2025.20160>
- Maluleke, A. F. (2025). AI adoption in African higher education: A systematic review of benefits and ethical implications. *Interdisciplinary Journal of Education Research*, 7(2), Article a05. <https://doi.org/10.38140/ijer-2025.vol7.2.05>
- Mithi, J., Madzvamuse, S., Mbanje, S., & Lomahoza, S. (2024, November 4–6). Generative artificial intelligence and formative assessment: Perspectives from higher education in South Africa [Conference session]. *Proceedings of the 1st International Conference on Education Research*, 1(1), 449–458. <https://doi.org/10.34190/icer.1.1.3231>
- Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open*, 5, Article 100151. <https://doi.org/10.1016/j.caeo.2023.100151>
- Ncube, P. D. N., Dzvapatsva, G. P., Matobobo, C., & Ranga, M. M. (2026). Redefining student assessment in AI-infused learning environments: A systematic review of challenges and strategies for academic integrity. *AI and Ethics*, 6, Article 68. <https://doi.org/10.1007/s43681-025-00871-w>

- Nicola-Richmond, K., Dawson, P., Partridge, H., & Macfarlane, S. (2025). It takes a village...: Program-wide approaches to redesigning assessment in a time of generative artificial intelligence (GenAI). *Journal of University Teaching and Learning Practice*, 22(7). <https://doi.org/10.53761/zpp2ja61>
- Page, E., Meyers, G., & Billings, E. (2024). Theory to practice: An assessment framework for generative AI. *Intersection: A Journal at the Intersection of Assessment and Learning*, 5(4), 114–126. <https://doi.org/10.61669/001c.124250>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice*, 20(2). <https://doi.org/10.53761/1.20.02.07>
- Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The Artificial Intelligence Assessment Scale (AIAS): A framework for ethical integration of generative AI in educational assessment. *Journal of University Teaching and Learning Practice*, 21(6), 49–66. <https://doi.org/10.53761/q3azde36>
- Perkins, M., Roe, J., & Furze, L. (2025). Reimagining the Artificial Intelligence Assessment Scale: A refined framework for educational assessment. *Journal of University Teaching and Learning Practice*, 22(7). <https://doi.org/10.53761/rrm4y757>
- Salinas-Navarro, D. E., Vilalta-Perdomo, E., Michel-Villarreal, R., & Montesinos, L. (2024). Using generative artificial intelligence tools to explain and enhance experiential learning for authentic assessment. *Education Sciences*, 14(1), Article 83. <https://doi.org/10.3390/educsci14010083>
- Wang, G., & Sun, F. (2025). A review of generative AI in digital education: Transforming learning, teaching, and assessment. *International Journal of Information and Communication Technology*, 26(19), 102–127. <https://doi.org/10.1504/IJICT.2025.146701>