

Process-Oriented Collaborative Essays: Empowering Undergraduate Students Through Assessment Redesign in the Age of Generative AI

Fahima Alam, Queen Mary University of London, United Kingdom
Jonathan Jackson, Queen Mary University of London, United Kingdom

The European Conference on Education 2026
Official Conference Proceedings

Abstract

The growing use of Generative Artificial Intelligence (GenAI) tools such as ChatGPT is reshaping how students approach learning and assessment in Higher Education (HE) as well as the viability of traditional assessment methods such as essays. While a significant amount of research has focused on student acceptance of GenAI through models such as the Technology Acceptance Model (TAM) and Unified Theory of Acceptance and Use of Technology (UTAUT), there is a pressing need for practical examples of how assessments in HE can remain valid and secure despite the prevalence of GenAI. This study explores the benefits and limitations of a collaborative, process-oriented assessment implemented with 23 first-year degree apprenticeship students at Queen Mary University of London. Student teams were required to engage with an iterative essay-writing task, producing multiple drafts and reflecting on how their work developed through collaboration with each other and optional use of GenAI tools to support their writing process. Drawing on Task-Technology Fit (TTF) and Institutional Theory, the study examines how GenAI aligns with academic writing practices and institutional expectations within an undergraduate assessment context. The findings suggest that GenAI can support the learning process through exploration of ideas and refinement of arguments over time, if appropriate assessment design scaffolding exists. The iterative and collaborative elements of the assessment appeared to support the development of students' academic voice, identity, and autonomy. The study highlights the value of designing assessments that prioritise process, reflection, and collaboration, enabling more meaningful and engaging learning experiences while strengthening academic integrity.

Keywords: GenAI, process-oriented assessment, assessment design, higher education

iafor

The International Academic Forum
www.iafor.org

Introduction

In recent years, the emergence of Generative Artificial Intelligence (GenAI) tools has disrupted the higher education (HE) landscape, with the conversation shifting from experimentation towards wider discussion on how these tools influence learning, teaching, and assessment (Cotton et al., 2024). Many arguments have centred on academic integrity, specifically the possibility of students submitting work generated in part or entirely by GenAI and the difficulties associated with detecting these practices (Kofinas et al., 2025; Perkins, 2023). Specific Large Language Model (LLM) chatbots such as ChatGPT have gained significant attention as they are able to generate ideas, propose structures, summarise information, refine arguments and draft text (Kasneci et al., 2023). Student use is also widespread with a 2026 study conducted by the Higher Education Policy Institute (HEPI) revealing that 94% of the students surveyed had used GenAI in some way for assessed work (Stephenson & Armstrong, 2026). Therefore, the issue is not about whether or how GenAI might be accepted. Rather, a more complex question must be asked: how should assessments be designed when emerging digital technologies can perform some of the required tasks that students have historically been expected to complete?

Bouaziz and Eddouada (2026) state that essay-based assessments are widely used as a method of evaluating the subject knowledge, critical thinking, and the analytical writing capabilities of the students. Nonetheless, the final essay only offers a partial representation or proxy of the intellectual processes that should be involved in essay writing, providing limited understanding on how students have built ideas, engaged with feedback, and revised their arguments over time (Lee, 2006; Mueller-Lyaskovets & Horner, 2021). As GenAI has become almost ubiquitous and capable of supporting and mimicking various aspects of academic writing, it is challenging to ascertain the extent to which a final written artefact reflects the knowledge, understanding, and intellectual development of the student (Ali & Maroulis, 2026). But this does not diminish the educational value of essay writing; instead, it raises the question of whether the final written essay, when considered in isolation, is able to offer sufficient evidence of the learning that an essay assessment is intended to measure. Cheng et al. (2024) assert that it is important to be attentive towards assessment approaches that make the reasoning, decision-making, and the development of the student more visible throughout the learning journey. It is with this principle in mind, that the work outlined in this paper was undertaken.

Significance of the Study

Many early studies around GenAI and assessment have emphasised detection, academic integrity, and the likelihood of students submitting work generated by GenAI as their own (Cotton et al., 2024; Perkins, 2023). As an emerging challenge and a wider pedagogical question, educators must ask what assessment approaches are most supportive of evidencing student learning, critical thinking and intellectual development (Ardito, 2025). Furze et al. (2024) argued that blocking or banning GenAI is unlikely to offer a sustainable response but instead illustrate how assessment might be redesigned across various levels of appropriate AI engagement. In a similar vein, Bannister et al. (2025) assert that it is important to assess the validity and susceptibility of current assessments within an environment where GenAI tools are widely available. Furthermore, systematic reviews of assessment within AI-infused learning environments have discovered the requirement to reconsider traditional assessment strategies to preserve meaningful assurance of learning (Al-Abdullatif & Alsubaie, 2024).

This paper presents one practical response to this challenge: a collaborative process-oriented essay assessment designed for first year undergraduate students. This paper also draws on two complementary theoretical perspectives: Task–Technology Fit (TTF) and Institutional Theory. TTF is used to examine the extent to which the capabilities of GenAI align with the requirements of specific academic writing tasks, while Institutional Theory provides a lens for considering whether such uses are consistent with institutional expectations surrounding academic integrity, authorship and legitimate evidence of student learning. Bringing these perspectives together facilitates the consideration of not only whether GenAI can effectively support an academic task, but also whether its use remains educationally and institutionally appropriate within the assessment context.

The core argument of this paper is that the value of process-oriented assessment does not lie in attempting to make assessment resistant to GenAI. Rather, it lies in making students' intellectual contribution more visible by providing evidence of how they develop ideas, exercise judgement, respond to feedback, collaborate with others and refine their work over time.

Literature Review

Impact of GenAI on Assessment in HE

Across the HE landscape, GenAI has become a prominent topic of discussion, arguably creating new opportunities for learning but at the same time, posing challenges for many extant approaches to assessment (Cotton et al., 2024). Kasneci et al. (2023) state LLMs can support many worthwhile activities such as personalised learning, idea generation, providing constructive feedback and other aspects of academic writing while acknowledging concerns around bias, misinformation, overdependence, and the extent to which students remain intellectually accountable for the work they generate. Hence, GenAI should not be comprehended as either inherently effective or detrimental to learning, but in the context of education it is important to understand how it could be employed in academic tasks and how students might continue to use it while maintaining critical judgement over generated outputs.

While there are legitimate reasons for some entirely AI-free learning environments, this has become an increasingly difficult proposition for HE to sustain. As noted in the introduction of this paper, as of 2026, 94% of students surveyed as part of HEPI research have used GenAI in some way for assessed work (Stephenson & Armstrong, 2026) so this cannot be treated lightly. As students are increasingly likely to encounter AI-supported practices within future professional environments, completely excluding GenAI from learning and assessment may restrict opportunities to develop the critical and responsible AI literacy required beyond higher education (Upsher et al., 2024; Su et al., 2026). Similarly, it is the responsibility of universities to ensure that assessment outcomes provide credible evidence of what students know and what they can do. This combination of competing demands makes GenAI a “wicked problem” for assessment (Corbin et al., 2025) in that there is no single correct solution to the problem. In some contexts, the prohibition of GenAI might protect certain conceptions of independent authorship but may minimise authenticity. On the other hand, unrestricted GenAI use might undermine individual students' ability to learn and be able to establish themselves in contemporary and future workplaces.

Thus, the challenge for HE extends beyond deciding whether GenAI should be prohibited or permitted. Fundamentally, a re-consideration is needed of how assessment can accommodate

increasingly AI-mediated learning while continuing to offer meaningful and valid evidence of knowledge, judgement, and intellectual development of the student.

Assessment Validity & Academic Integrity in the Age of GenAI

Within HE, GenAI has sparked two important concerns: assessment validity and academic integrity. From the viewpoint of academic integrity, attention has been concentrated on whether students are the main authors of an essay, for example, and whether the usage of GenAI has been consistent with the expectations of the institution in the context of appropriate academic practice (Cotton et al., 2024; Kofinas et al., 2025). Essentially, has the student remained intellectually accountable for their work or have they offloaded it to an extent which is deemed academic misconduct?

On the other hand, assessment validity raises different concerns such as whether the assessment facilitates the production of suitable evidence by students that they have the knowledge and skills that the assessment was anticipated to measure. Of course, academic integrity and assessment validity can overlap but they should not be conflated. For instance, a student can make use of GenAI within the institutional boundaries and therefore satisfy the academic integrity requirements, but the resulting assessment might offer limited suitable evidence of a particular capability or competency.

Importantly, the use of GenAI does not negate the assessment validity where students are still accountable for assessing the output, making decisions, and building their own argument. This is validated by two studies which emphasise the significance of considering the relationship between GenAI usage and what the assessment is intended to evidence (Bannister et al., 2025; Sharma, 2026).

In fact, Sharma (2026) suggests that the student's judgement can be more visible when the assessment is supported by GenAI. The quality of the final submission is significant, but this can be reinforced by supplementary evidence that shows how the student achieved the outcome. Essentially, process-oriented assessment can help in understanding the reasoning, decisions, and development of the students' knowledge and at the same time evaluate the final written work.

Process-Oriented Assessments

Process-oriented assessment offers a possible response to the various challenges surrounding opaque written assessments in the age of GenAI. Conventionally, an essay assessment is focused on the quality of the final submission where the final written essay acts as the sole piece of evidence for evaluating the performance of the student (Lee, 2006). Process-based writing literature has argued that writing should be comprehended as an activity of development that involves drafting, planning, feedback, revision, and reflection instead of solely focusing on the finished product (Lee, 2006; Mueller-Lyaskovets & Horner, 2021). Thus, a process-oriented assessment helps in broadening the evidence that is available by making the chosen stages of a student's development more explicit.

There is an important distinction between accepting a suggestion generated by GenAI without questioning it and evaluating the suggestion against academic evidence, discussing it with others, and then deciding whether to reject or accept it. Cheng et al., (2024) conducted research examining human-AI collaborative writing and they similarly recommend that process

evidence can offer information about how students are interacting with GenAI that might not be visible from the completed text alone.

Nonetheless, simply asking students to provide several drafts of their work does not necessarily offer strong evidence of learning. It is important for the process-oriented assessments to emphasise the purpose and quality of development instead of the number of documented stages (Lee, 2006). All too easily, an assessment designed with good intentions could lead to students making relatively trivial changes between drafts simply to satisfy the requirement of showing multiple revisions, or they could concentrate on performatively creating process-oriented artefacts rather than engaging in the process in a way that develops their thinking. The value of process evidence therefore lies less in showing that revisions took place (which risks a performative response) and more in revealing the judgement behind those revisions. For example, students might explain why an argument was reconsidered, how evidence changed their position, how peer feedback influenced a decision, or why a GenAI-generated suggestion was accepted, modified or rejected.

This approach also connects with emerging discussions of authentic assessment in relation to the pervasive presence of GenAI. Su et al. (2026) argue that assessment needs to recognise the evolving role of GenAI while continuing to create opportunities for students to demonstrate meaningful learning. Similarly, frameworks for responsible GenAI use emphasise that different levels of AI involvement should be considered in relation to the purpose and intended learning outcomes of an assessment (Ucan & Demirel Ucan, 2026). The transition away from product-oriented assessment towards process-oriented assessment should therefore not be understood as abandoning the final essay. Instead, the final artefact can remain an important indicator of achievement while being complemented by evidence of how students developed it (Midford, 2025). In this way, iteration, collaboration and reflection can provide additional insight into students' reasoning, decision-making and intellectual contribution (Cheng et al., 2024).

Collaboration, Reflection and Academic Voice

Collaboration is seen to be pedagogically important as it facilitates important human-to-human interaction with the intention of mitigating risk of student isolation. Also, collaborative writing requires students to respond to different interpretations, negotiate meaning, justify choices and coordinate the development of a shared text. Scaffolding these activities in class provides visible social learning and can offer a supplementary source of responsibility as decisions are exposed to peers instead of working privately (Gündüz, 2023; Khuder & Negretti, 2026).

Additionally, reflection can provide another layer to the assessment. The concept of reflection-on-action (Schon, 1983) is useful for helping students understand how they might retrospectively assess decisions and experiences. Within an assessment which utilises GenAI, reflective documentation can help in revealing not only what students have changed but why they have changed them. For example, providing an explanation for why an AI generated suggestion was rejected offers evidence of evaluative judgement that would be absent from the final artefact. Therefore, reflection can operate simultaneously as a learning activity as well as evidence for process-oriented assessment.

This paper proposes that the combination of collaboration and reflection together are particularly valuable within the scope of process-oriented assessment. As mentioned previously, collaboration can create opportunities for ideas to be negotiated, questioned, and refined through interaction with others whereas reflection can offer students to explain the

reasoning behind the decisions. These approaches transition responsibility away from relatively ineffective declarations of GenAI use towards demonstrating how students exercise judgement over its use. This approach seems to be consistent with emerging assessment frameworks encompassing GenAI that suggest aligning the permitted usage of AI with intended learning outcomes and at the same time retaining meaningful opportunities for critical thinking and human input (Furze et al., 2024; Perkins et al., 2024).

Theoretical Frameworks

Technology-Task Fit (TTF)

TTF considers whether the capabilities of a technology align with the requirements of a particular task. In the context of this study, it is helpful to examine whether GenAI genuinely supports different aspects of academic writing or whether the fit is more limited. In their review of TTF literature, Spies et al. (2020) showed that the TTF framework has been applied across a wide range of technology-use contexts. In relation to GenAI, TTF is useful because it moves the discussion away from simply asking whether AI is beneficial or harmful to academic writing. Instead, it allows us to ask a more specific question: how well do the capabilities of GenAI fit different academic writing tasks?

More recent research further supports the relevance of TTF to GenAI. Wang et al. (2026) examined the relationship between TTF, perceived usefulness and responsible GenAI use among 280 university students. They found that a stronger fit between academic tasks and GenAI capabilities was associated with greater perceived usefulness and more responsible use. The study also extends established technology-adoption approaches, such as the Technology Acceptance Model (TAM), by considering not only whether students find a technology useful, but also how responsibly it is used. This is particularly relevant to assessment, where the concern is not simply whether students use GenAI, but how its use relates to the intended learning outcomes.

In academic writing, for example, GenAI may have a strong task fit for generating possible counterarguments, exploring alternative structures, brainstorming ideas or identifying language issues. However, the fit becomes more complex when students are expected to demonstrate their own judgement, evaluate the quality of evidence, apply professional experience or negotiate different viewpoints with peers. TTF therefore allows GenAI to be considered in relation to specific academic writing activities, rather than treating the essay as one single task.

Institutional Theory

Institutional Theory helps in the understanding of expectations and contextual norms within which universities operate. Assessment does not take place in isolation; it is influenced by institutional policies, regulatory requirements and established expectations about what counts as evidence of student learning (Cai & Mehari, 2015). This means that assessment design is not only shaped by what works well for teaching and learning. Universities must also consider issues such as authorship, fairness, academic integrity, quality assurance and the credibility of the qualifications they award.

Using Institutional Theory alongside TTF facilitates a consideration of both task fit and institutional fit. For example, GenAI may be very effective at generating or improving written text, meaning that the technology can fit certain academic writing tasks well. However, this

does not necessarily mean that its use is appropriate for every assessment. If learning outcomes require students to develop, evaluate and communicate their own argument, extensive reliance on GenAI may make it more difficult to interpret the student's work without any evidence of their process. At the same time, allowing appropriate use of GenAI may be valuable where developing AI literacy is relevant to the learning outcomes and students' future professional practice. Therefore, the question is not simply whether GenAI can support a particular task, but whether the way it is used also fits the purpose of the assessment and the expectations of the institution.

Methodology

Rather than a controlled research intervention, the work outlined in this paper was undertaken as part of programme-wide pedagogical redesign. This study therefore adopts a practice based exploratory perspective and focusses on the resulting observations as emerging insights instead of generalisable empirical findings.

The collaborative process-oriented essay assessment was designed for first year Business Analyst degree apprenticeship students at Queen Mary University of London. The cohort size was 23 students, made up of 11 self-selecting teams, one of which was a team of 3.

Students were required to develop a 1500-word essay in their teams and document at least two stages of iterative development. Each team was required to choose from three pre-determined essay titles, which related to the main topic of the module. The assessment criteria made clear how evidence of iteration and collaboration would contribute to the final grade, as demonstrated by Table 1. While by no means perfect, these assessment criteria ensured that the assessment was structurally designed to require evidence of these activities (Corbin, Dawson, et al., 2025).

Students were permitted to GenAI within clear boundaries, but use of GenAI was not mandatory. The following Generative AI guidance was included in the assessment brief for clarity and to encourage good academic practice and avoid common pitfalls:

Use of Generative AI (e.g., MS Copilot or ChatGPT) is permitted for this assessment but must not be used to generate text which you then submit as your own as this is academic misconduct. If using spelling or grammar assistive tools, avoid using any Generative AI functionality which rewrites non-trivial portions of your work. Avoid relying on Generative AI to format your references as these tools can often produce references which do not exist which will significantly undermine the credibility of your work and may be considered academic misconduct.

While this statement can be viewed as discursive rather than structural, it was still deemed necessary for the purposes of making the expectations of the assessment clear and to provide reasons for how inappropriate use of GenAI can lead to poor outcomes.

Table 1*Sample Assessment Criteria Related to Collaboration and Iteration*

Criteria	0-29 (F)	30-39 (E)	40-49 (D)	50-59 (C)	60-69 (B)	70-79 (A)	80-100 (A+)
Collaborative Contribution & Knowledge Co-construction (20%)	No meaningful evidence of collaboration.	Minimal collaboration; weak evidence of joint working.	Limited evidence of collaboration; engagement appears inconsistent.	Basic collaboration evident but contributions may be uneven.	Effective collaboration with some evidence of shared contribution.	Very effective collaboration with clear evidence of peer engagement and joint responsibility.	Exceptional evidence of balanced collaboration, shared decision-making, and collective knowledge construction throughout the task.
Iterative Development & Evidence of Revision (20%)	No evidence of iterative development.	Drafts submitted but show negligible revision or learning progression.	Minimal evidence of iteration; little meaningful development between drafts.	Some iteration evident; changes are mostly surface level with minimal explanation.	Evidence of at least 2 iterations with noticeable improvement, though justifications may lack depth.	Strong evidence of at least 2 meaningful iterations with clear improvement and rationale for revisions.	In addition to A-grade requirement: Evidence of insightful reflections on the collaborative process and shared learning.

Discussion

This section outlines some of the co-authors' insights and reflections after administering the process-oriented collaborative essay assessment for the first time in early 2026.

GenAI as a Source for Interrogation Instead of Offloading

One noteworthy feature that was evident in some of the submissions was that students who used GenAI queried the suggestions instead of simply accepting them. For example, in some cases the interaction of counterarguments was followed by revisions, active challenging, or rejection. This pattern is vital as it helps in demonstrating that usage of GenAI can indeed incorporate and even stimulate intellectual work instead of replacing it. Therefore, AI can be seen as a dialogic resource that helps stimulate further thought (Miao & Holmes, 2023).

Within the context of TTF, this represents an alignment between task and technology where GenAI is appropriate in rapidly producing alternate positions, however the educational value is based on whether students are being held accountable and sufficiently motivated to assess those positions. In a similar vein, technological capabilities can help in supporting learning or they can undermine it and this is based on the subsequent activities of the student. Therefore, the assessment design should emphasise not only whether GenAI is permitted but it should

offer structural requirements on how positive interactions with GenAI tools can be sufficiently evidenced.

Collaborative Learning as Process Evidence

The 11 student teams displayed various approaches to collaboration during the module as they worked on their assessment. Students shared their drafts, exchanged peer feedback, and revised the work in an iterative manner. Some also leveraged Microsoft Teams video call and chat discussions that offered additional opportunities for them to informally present and explain the reasoning behind specific decisions. Therefore, collaboration offered a social dimension in the process-oriented assessment. All these aspects are important as collaborative writing is not simply about dividing a task between team members. This is evident in recent research that illustrates effective collaboration can involve negotiation, collective decision-making, shared goal setting, and regulation of the developing text (Khuder & Negretti, 2026).

Khuder and Negretti (2026) discovered that students who developed socially shared regulation discussed issues collectively, negotiated various feedback and made decisions together. Additionally, they further advocate that iterative processes are useful as students can justify decisions and can offer clarification which can help in strengthening the collaborative regulation. Therefore, this offers a significant insight for the current study where students provided an explanation to their peers, responded to disagreements, justified their revisions, and offered explanations that may be less visible in individual writing.

In the current era of GenAI ubiquity, the collaborative dimension is important. Some recent work that has integrated peer feedback and GenAI into academic writing recommends that human peer interaction can help in complementing writing that is supported by AI by offering opportunities for critical engagement (Orvalho & Monteiro, 2025). Thus, collaboration can offer an extra layer of analysis around the development where students are required to explain and negotiate choices instead of making them in isolation.

Even though collaboration can offer opportunities, it should not be treated as singular evidence of equitable learning. Likewise, working in a group does not necessarily mean that meaningful collaboration takes place. According to Khuder and Negretti (2026), collaboration is stronger when students can share their goals, discuss decisions, and take collective accountability for the work. This is significant to the current study where 11 teams demonstrated various approaches to collaboration and revision, some more effective than others. Consequently, the authors note that the quality of collaboration could be improved with more scaffolded collaborative activities in class to help students experience different ways of collaborating so that they can then leverage these approaches more autonomously for their collaborative assessment.

Limitations of Process-Oriented Documentation

Some students informally mentioned that the required process-based documentation seemed like a lengthy and time-consuming process. Also, some students were not initially sure how to present a meaningful commentary on their process. These lessons are important to consider as it demonstrates how the purpose behind the process-based evidence needs to be made clear to students and that more examples may be beneficial, particularly for students relatively new to university-level education.

Another limitation of process-oriented assessment is the risk of “empty performativity” where apparent evidence produced by students does not reflect any genuine intellectual development. When students understand that collaboration, revision, and reflection are being assessed, they may lean towards producing evidence of activities that meet the perceived expectations of the marker. In a systematic review by Ross et al., (2024) on reflective assessment, it was discovered the expectations of the assessment can encourage performative instead of genuine reflection. Hence, reflections, notes, and records of collaboration should not automatically be viewed as satisfactory evidence of learning.

Conclusion

The response to GenAI by HE must move beyond attempts to detect whether students have used GenAI tools as part of an assessment. In this paper, collaborative and process-oriented assessment is shown to afford opportunities to make students’ learning visible through peer interaction, iterative drafting, revision and reflection. All these elements contribute to a multi-faceted body of evidence which can demonstrate how students are developing arguments, building judgement during the writing process, and responding to feedback. Nonetheless, process-oriented assessments should not be viewed as a solution suitable for all contexts. The differences in participation across groups and the additional workload (perceived or actual) related to documenting an iterative process are factors which can still undermine the validity of the assessment without careful scaffolding and student support.

The authors of this paper plan to strengthen future assessments by communicating more clearly to students the purposefulness of evidencing their process while minimising less valuable documentation. Other emerging frameworks will also be explored such as the Voice-First Written Assessment (VFWA) approach proposed by Webb-Davies et al. (2026) which may offer a beneficial extension to process-oriented essay assessments. VFWA involves capturing “evidence of [students’] task-related reasoning under observed conditions before generating a final written submission” in an environment where GenAI or other assistive digital technologies are available. The approach looks to establish a clear association between the final written work and the reasoning of the student instead of attempting to eliminate the use of technology in writing, which aligns with the underlying principles of the work outlined in this paper. Future research would be able to examine whether the combination of these approaches leads to improved quality of evidence of student judgement while remaining manageable for students and those assessing the work.

In summary, the work presented in this paper suggests that a combination of scaffolded collaborative work and process-oriented assessment design can lead to conditions where reasoning, judgement, and iterative development can remain more visible and offer valuable evidence upon which evaluation of students’ work can be based.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors declare that ChatGPT was used in refining some language used in the first draft of the manuscript. The usage was limited to correcting grammatical and spelling errors and rephrasing statements for accuracy and clarity. Subsequently, a full human review and edit was undertaken prior to submission of the manuscript for consideration in the conference proceedings. The authors further declare that, apart from ChatGPT, no other AI or AI-assisted technologies have been used to generate content in writing the manuscript. The ideas, design,

procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of practice-based research.

References

- Al-Abdullatif, A. M., & Alsubaie, M. A. (2024). ChatGPT in Learning: Assessing Students' Use Intentions through the Lens of Perceived Value and the Influence of AI Literacy. *Behavioral Sciences*, 14(9), 845. <https://doi.org/10.3390/bs14090845>
- Ali, R., & Maroulis, J. (2026). Assessment validity in the age of generative artificial intelligence: A critical review. *Assessment & Evaluation in Higher Education*, 0(0), 1–16. <https://doi.org/10.1080/02602938.2026.2683046>
- Ardito, C. G. (2025). Generative AI detection in higher education assessments. *New Directions for Teaching and Learning*, 2025(182), 11–28. <https://doi.org/10.1002/tl.20624>
- Bannister, P., Santamaría Urbieto, A., & Brufau Alvira, N. (2025). Appraising higher education assessment validity: Development of the PANDORA GenAI Susceptibility Rubric. *Journal of Applied Learning & Teaching*, 8(1). <https://doi.org/10.37074/jalt.2025.8.1.20>
- Bouaziz, K., & Eddouada, S. (2026). AI-Assisted Academic Writing and Publishing in Morocco: Doctoral Students' Ethical Decision-Making Accounts. *Arab World English Journal*, (3), 268–286. <https://doi.org/10.24093/awej/A3.17>
- Cai, Y., & Mehari, Y. (2015). The Use of Institutional Theory in Higher Education Research. In J. Huisman & M. Tight (Eds), *Theory and Method in Higher Education Research* (Vol. 1, pp. 1–25). Emerald Group Publishing Limited. <https://doi.org/10.1108/S2056-375220150000001001>
- Cheng, Y., Lyons, K., Chen, G., Gašević, D., & Swiecki, Z. (2024). Evidence-centered Assessment for Writing with Generative AI. *Proceedings of the 14th Learning Analytics and Knowledge Conference*, 178–188. <https://doi.org/10.1145/3636555.3636866>
- Corbin, T., Bearman, M., Boud, D., & Dawson, P. (2025). The wicked problem of AI and assessment. *Assessment & Evaluation in Higher Education*, 51(4), 736–752. <https://doi.org/10.1080/02602938.2025.2553340>
- Corbin, T., Dawson, P., & Liu, D. (2025). Talk is cheap: Why structural assessment changes are needed for a time of GenAI. *Assessment & Evaluation in Higher Education*, 50(7), 1087–1097. <https://doi.org/10.1080/02602938.2025.2503964>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (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>
- Furze, L., Perkins, M., Roe, J., & MacVaugh, J. (2024). The AI Assessment Scale (AIAS) in action: A pilot implementation of GenAI-supported assessment. *Australasian Journal of Educational Technology*. <https://doi.org/10.14742/ajet.9434>

- Gündüz, Z. E. (2023). Teaching Writing with Wiki-Based Collaborative Writing Tasks in an EFL Context at Higher Education. *Acuity: Journal of English Language Pedagogy, Literature and Culture*, 8(1), 68–86.
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Khuder, B., & Negretti, R. (2026). Collaborative writing regulation: A comparative case study of co-regulation and socially shared regulation in higher education. *Higher Education*, 91(2), 677–699. <https://doi.org/10.1007/s10734-025-01437-9>
- Kofinas, A. K., Tsay, C. H., & Pike, D. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology*, 56(6), 2522–2549. <https://doi.org/10.1111/bjet.13585>
- Lee, Y.-J. (2006). The process-oriented ESL writing assessment: Promises and challenges. *Journal of Second Language Writing*, 15(4), 307–330. <https://doi.org/10.1016/j.jslw.2006.09.003>
- Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. <https://doi.org/10.54675/EWZM9535>
- Midford, S. (2025). Is the essay history? Rethinking effective assessment in the age of generative AI. *History Australia*, 22(4), 563–567. <https://doi.org/10.1080/14490854.2025.2570457>
- Mueller-Lyaskovets, T., & Horner, O. (2021). Integrating Formative Assessment with Foreign Language (English) Process Writing Instruction: Lessons from Two College Writing and Reading Classes in Germany. *Journal of Academic Writing*, 11(1), 62–79. <https://doi.org/10.18552/joaw.v11i1.499>
- Orvalho, L., & Monteiro, A. F. (2025). *IA Generativa e Feedback entre Pares na Escrita Acadêmica* [Generative AI and peer feedback in collaborative academic writing]. *Saber e Educar*, 34(1). <https://doi.org/10.25767/SE.V34I1.42564>
- 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), e635–e635. <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(06). <https://doi.org/10.53761/q3azde36>

- Ross, M., Bohlmann, J., & Marren, A. (2024). Reflective writing as summative assessment in higher education: A systematic review. *Journal of Perspectives in Applied Academic Practice*, 12(1). <https://doi.org/10.56433/jpaap.v12i1.577>
- Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.
- Sharma, S. (2026). Educational integrity in GenAI-augmented assessment: Making judgement visible. *International Journal for Educational Integrity*, 22(1), 6. <https://doi.org/10.1007/s40979-026-00216-y>
- Spies, R., Grobbelaar, S., & Botha, A. (2020). A Scoping Review of the Application of the Task-Technology Fit Theory. In M. Hattingh, M. Matthee, H. Smuts, I. Pappas, Y. K. Dwivedi, & M. Mäntymäki (Eds), *Responsible Design, Implementation and Use of Information and Communication Technology* (Vol. 12066, pp. 397–408). Springer International Publishing. https://doi.org/10.1007/978-3-030-44999-5_33
- Stephenson, R., & Armstrong, C. (2026). *Student Generative Artificial Intelligence Survey 2026* (No. 199). HEPI. <https://www.hepi.ac.uk/reports/student-generative-ai-survey-2026/>
- Su, F., O'Dea, M., Wood, M., & Seidl, E. (2026). Embracing a new world: Authentic assessment designs in an age of generative artificial intelligence (GenAI). *Perspectives: Policy and Practice in Higher Education*, 1–12. <https://doi.org/10.1080/13603108.2025.2601741>
- Ucan, S., & Demirel Ucan, A. (2026). Generative AI in Student Writing: A Six-Level Taxonomy for Responsible Use in Higher Education Assessment. *Journal of Academic Ethics*, 24(2), 63. <https://doi.org/10.1007/s10805-026-09739-y>
- Upsher, R., Heard, C., Yalcintas, S., Pearson, J., & Findon, J. (2024). Embracing generative AI in authentic assessment: Challenges, ethics, and opportunities. In *Using Generative AI Effectively in Higher Education*. Routledge.
- Wang, Q., Wang, Y., & Amini, M. (2026). Associations of task technology fit and perceived usefulness with responsible generative AI use among university students. *Scientific Reports*, 16(1), 2759. <https://doi.org/10.1038/s41598-025-32692-6>
- Webb-Davies, K., Ratner, S., & Bassett, M. A. (2026). Voice-First Written Assessment (VFWA): Restoring evidence of reasoning through structural redesign (Tyh8z_v1). EdArXiv. https://osf.io/preprints/edarxiv/tyh8z_v1/