Navigating GenAI in Psychology Education: Assessment Validity, Academic Integrity, and the Realities of Teaching in an AI-Rich Era

Jean-François Delvenne[1] 📍

Abstract

This review examines how psychology educators are responding to the rapid rise of generative artificial intelligence (GenAI), focusing on implications for assessment validity, academic integrity and organisational learning. Although universities have issued policy guidance, these frameworks often overlook psychology's distinctive epistemic, methodological and pedagogical practices. Drawing on empirical research, sector reports, survey findings and my experience as academic integrity lead in a UK university, the review identifies five interconnected challenges: unreliable detection technologies, ambiguity in marking and feedback, threats to the validity of psychology-specific assessments, increasingly complex integrity casework, and limited institutional support. It argues that psychology is well positioned to provide sector-wide leadership because of its emphasis on empirical reasoning, ethical judgement and reflective practice. The article synthesises emerging discipline-sensitive strategies and offers a forward-looking agenda for research, assessment design and staff development, emphasising approaches that foreground reasoning processes, ethical awareness and critical engagement with GenAI. It concludes by calling for a coordinated institutional response that integrates clearer policy, systematic staff training and strengthened communities of practice to support academic integrity and student learning in a GenAI-rich environment.

Keywords

Generative AI, academic integrity, assessment validity, psychology education, ethics, GenAI literacy, higher education policy

[1]School of Psychology, University of Leeds, Leeds, UK

Corresponding Author:
Jean-Francois Delvenne, School of Psychology, University of Leeds, Leeds, LS2 9JT, UK.
Email: J.f.delvenne@leeds.ac.uk.

Introduction

Psychology departments across the higher education sector are experiencing strong and immediate pressures arising from the widespread use of generative artificial intelligence (GenAI) tools. While these pressures are shared with other disciplines, psychology faces a distinctive constellation of challenges because its core learning outcomes depend heavily on empirical reasoning, ethical judgement, methodological understanding and reflective practice. As the academic integrity lead in a university psychology department, responsible for reviewing evidence, advising colleagues, interpreting policy and ensuring fair decision-making, I have observed how GenAI is disrupting long-established assumptions about authorship, learning and assessment. These disruptions are most visible in coursework formats that constitute the backbone of psychology education. Laboratory reports require students to demonstrate understanding of design principles, operational definitions, sampling decisions, ethical considerations and statistical interpretation. Critical evaluations require detailed engagement with empirical literature, methodological critique and evidence-based argumentation. Reflective assignments invite students to demonstrate metacognitive insight, personal development, and ethical awareness. GenAI can produce superficially polished versions of these outputs, but the reasoning processes that underpin them, the very constructs the assessments are intended to measure, may be absent, concealed or distorted.

Research indicates that students frequently use GenAI not only for legitimate support such as proofreading but also for summarising empirical literature, structuring essays or generating first drafts (Black & Tomlinson, 2025; Darvishi et al., 2024). While students may view these practices as efficient or benign, they raise significant questions about authorship, transparency, fairness and learning. In psychology, uncritical use risks obscuring weaknesses in methodological understanding, ethical reasoning and information literacy, especially when AI-generated content appears fluent yet lacks nuance or conceptual depth. Recent findings suggest that procedural reliance on GenAI is associated with lower learning outcomes (Pallant et al., 2025), reinforcing concerns that overuse may impair disciplinary skill development. Although sector bodies such as Jisc (2023), QAA (2024) and UCISA (2025) have highlighted the disruptive implications of GenAI, discipline-specific insights remain limited. Psychology educators report uncertainty about interpreting polished but shallow writing, preserving the validity of assessments designed to elicit reasoning, and managing rising integrity casework. Students, too, perceive psychology as particularly vulnerable to GenAI misuse (Acosta-Enriquez et al., 2024; Tierney et al., 2025).

Taken together, these challenges highlight the need for a theoretically grounded account of how GenAI interacts with disciplinary practice. To provide such a framework, this review draws on three conceptual anchors – assessment validity, academic integrity and organisational learning – which together provide a coherent framework for understanding how GenAI reshapes psychological education. These anchors illuminate how GenAI disrupts not only specific assessments or marking processes but the wider ecology of teaching, learning and quality assurance within psychology. The review identifies five key challenges, namely: (1) unreliable detection, (2) marking and feedback uncertainty, (3) threats to validity, (4) burdens on integrity processes and (5) insufficient institutional support. Rather than merely cataloguing problems, the review uses these frameworks to provide a structured account of emerging disciplinary responses and to outline a constructive path forward for psychology education.

Challenges in Assessment and Integrity Practices: From Detection to Casework

The five challenges summarised in Table 1 are analytically distinct but practically interdependent: problems in detection feed into marking uncertainty; threats to validity exacerbate integrity

Table 1. Key Challenges, Their Psychology-Specific Manifestations and Emerging Responses.

Challenge	Psychology-Specific Illustration	Emerging Response / Example Strategy
1. Unreliable detection	Highly polished reflective work lacking personal insight; unrealistically tidy statistical reporting	Scaffolded drafting; oral verification; reflective justification of analytic and methodological decisions
2. Marking & feedback uncertainty	Fluent prose masking weak reasoning; doubts about whether comments supports actual learning	Feedback emphasising reasoning processes; in-class clarification and diagnostic tasks
3. Threats to assessment validity	AI-generated APA lab reports; fabricated datasets; superficial critiques of empirical work	Mixed-format assessments; process-based tasks; critical evaluation of GenAI outputs
4. Integrity casework burden	Ambiguous paraphrasing; work with no drafts; unclear evidence of authorship	Standardised evidence logs; oral follow-ups; discipline-specific case exemplars
5. Limited institutional support	Generic policy frameworks lacking relevance to psychology-specific tasks	Communities of practice; targeted GenAI training; discipline-sensitive exemplars

Note. AI=artificial intelligence; GenAI= generative AI.

casework; and limited institutional guidance compounds all of these issues. Table 1 highlights how each challenge manifests in psychology-specific contexts and summarises emerging responses.

The Detection Dilemma

Although educators across disciplines struggle to identify GenAI-generated content, psychology faces particular difficulties due to the nature of its assessments. Reflective writing, ethical analyses and empirical reports often reveal problems not through direct evidence of misconduct but through stylistic or conceptual incongruities: impersonal tone in reflective submissions, statistically implausible results sections or generic methodological commentary that appears detached from specific task requirements. Markers frequently describe a sense that 'something is not right' without being able to articulate or evidence the concern. This interpretive uncertainty is well documented (Kofinas et al., 2025), and colleagues often seek guidance on whether stylistic oddities merit escalation.

Detection tools such as Turnitin or GPTZero have been widely adopted, yet remain unreliable, unvalidated and prone to false positives, especially for non-native English speakers or students producing formulaic text (Chechitelli, 2023; Perkins, Roe, et al., 2024; Yeadon et al., 2023). Equally problematic are false negatives, where GenAI has been used lightly, such as for paraphrasing or structural editing, yet remains undetectable. These technologies cannot confirm GenAI authorship, as the QAA (2024) has emphasised, and their limitations leave staff uncertain about when and how to act. This ambiguity alters staff behaviour in ways that directly reshape assessment practice. Some markers become hesitant to raise cases without conclusive evidence; others adopt more defensive marking practices. The cumulative effect is a shift in the epistemic foundations of assessment: trust, once central to the marker–student relationship, becomes fragile and educators describe the emotional strain of feeling responsible for identifying unacknowledged GenAI use despite inadequate tools or support.

Marking and Feedback Uncertainty

When markers cannot be confident that submitted work reflects students' reasoning, the feedback process becomes compromised. Psychology's feedback culture relies heavily on diagnosing misconceptions, guiding analytical development, and supporting reflective learning (Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006). Yet if the prose appears unusually fluent, overly structured or inconsistent with earlier work, markers may question the authenticity of the submission and, by extension, the value of offering detailed developmental commentary. These concerns echo wider sector observations that GenAI obscures authorship and disrupts markers' interpretive confidence (Abdelaal & Al Sawy, 2024; Bobula, 2024).

This uncertainty is particularly acute in psychology because assessments are designed to reveal reasoning processes. A flawlessly written lab report that misinterprets analysis of variance (ANOVA) output raises the question of whether the misunderstanding lies with the student or with GenAI-generated text. Reflective assignments with generic or decontextualised insights raise similar doubts about the provenance of ideas. Faced with this ambiguity, some markers simplify feedback, shift towards generic comments, or avoid offering detailed developmental guidance. Such adaptations weaken the formative function of feedback and undermine students' development of feedback literacy – the capacity to understand, interpret and act on comments (Carless & Boud, 2018). If feedback no longer maps reliably onto student reasoning, its pedagogical value diminishes. Psychology educators describe frustration and a sense of futility, uncertain whether feedback supports authentic learning or merely assists students in refining future GenAI use. This erosion of trust threatens one of the discipline's key pedagogical strengths: the use of feedback to scaffold the development of scientific reasoning, ethical judgement and reflective capacity.

Threats to Assessment Validity

While marking uncertainty affects day-to-day teaching practice, its implications for assessment validity are profound. Validity theory emphasises the need for alignment between intended learning outcomes, assessment tasks, and the constructs they measure (Biggs, 1996). In psychology, written assessments are intended to elicit students' reasoning processes, methodological understanding, and critical engagement with evidence. GenAI undermines these assumptions by enabling students to generate plausible text without demonstrating the targeted cognitive processes, an issue increasingly noted across higher education (Abdelaal & Al Sawy, 2024; Bobula, 2024).

Laboratory reports illustrate this problem vividly: GenAI can generate APA-style writing, propose hypotheses and even simulate statistical outputs, yet these components may contain conceptual errors, fabricated data or inappropriate methodological choices. Literature reviews generated with AI tools often lack methodological nuance, misrepresent research findings or present fabricated citations. Research-design tasks can be similarly compromised when GenAI templates supersede students' own reasoning.

Empirical evidence amplifies these concerns. AI-generated exam responses have been shown to match or exceed the quality of human-written answers (Scarfe et al., 2024), and GenAI tools can perform competitively on Multiple Choice Question assessments (Newton & Xiromeriti, 2024). Together, these findings heighten doubts about whether traditional coursework formats can still differentiate between superficial fluency and deep understanding.

From a validity perspective, these developments create misalignment between intended learning outcomes and the evidence used to judge student achievement. Instead of capturing students'

conceptual reasoning, assessments may now measure their ability to utilise GenAI effectively. Without intervention, the evidentiary value of key psychology assessments – lab reports, critiques and reflections – risks erosion. Departments are experimenting with process-focused alternatives such as scaffolded drafting, annotated bibliographies and oral components (Cotton et al., 2024; Francis et al., 2025; Lee et al., 2024; Malik et al., 2025; Tierney et al., 2025), but these innovations require systematic institutional support and time for evaluation.

Integrity Casework and Procedural Burden

As assessments become increasingly vulnerable to GenAI, integrity casework grows more complex. Unlike plagiarism, which involves identifiable sources, GenAI leaves no external trace. Students may use GenAI for drafting, paraphrasing, structuring or proofreading without malicious intent. Without drafts, version histories or explicit disclosures, even well-founded suspicions rarely meet evidentiary thresholds, a challenge widely recognised across the sector (Abdelaal & Al Sawy, 2024). Departments report increased case volumes, longer decision timelines and significant variation in case handling (Cotton et al., 2024; Slimi, 2023). Psychology educators describe the emotional and cognitive load associated with navigating ambiguous cases, supporting anxious students and interpreting policy in the absence of reliable detection tools. These patterns mirror findings elsewhere in higher education, where staff report heightened uncertainty and procedural strain arising from GenAI-related misconduct concerns (Tierney et al., 2025). Integrity leads face particular pressure as they mediate between institutional expectations, student welfare and staff concerns. From an organisational-learning perspective, these pressures are exacerbated by the absence of shared case exemplars, feedback loops, or cross-departmental learning structures. As a result, educators often work in isolation, improvising responses that lead to inconsistency and frustration. Strengthening these organisational mechanisms is essential if institutions are to respond coherently and support staff effectively.

Institutional Responses

Limitations of Current Institutional Guidance

Universities have responded to GenAI by issuing policy updates, staff briefings and guidance for students (e.g., Russell Group, 2023). These frameworks typically designate acceptable uses (e.g., idea generation and proofreading) and prohibited uses (e.g., producing assessable content). Yet such categorisations often lack relevance for psychology-specific tasks requiring personal reflection, ethical reasoning, methodological justification or empirical interpretation. Allowing GenAI for 'structuring', for instance, offers little clarity for educators evaluating impersonal reflective submissions. Policies relying on student disclosure, such as traffic-light systems or the AI Assessment Scale (Perkins, Furze, et al., 2024), are limited by unverifiability (Corbin et al., 2025). Studies show that institutional policies are frequently generic, inconsistently implemented and lacking in practical disciplinary guidance (An et al., 2025; McDonald et al., 2025). Moreover, guidance often arrives after module approval deadlines, constraining educators' ability to redesign assessments or adjust marking procedures. These shortcomings contribute to reactive rather than strategic institutional responses. As one colleague reflected, 'We are being asked to reinvent assessment on the fly', a sentiment that underscores the need for discipline-sensitive policies and structured staff development, themes elaborated in the 'Clarifying Policy and Supporting Fair Practice in Psychology' and the 'Supporting Psychology Educators: Training and Peer Learning' sections.

Evolving Assessment Strategies in Psychology

Assessment redesign is one of the most active and consequential areas of adaptation within psychology departments, and one that sits at the intersection of assessment validity, academic integrity and disciplinary pedagogy. Psychology educators are acutely aware that many traditional assessment formats – laboratory reports, critical evaluations, reflective assignments and research proposals – were developed for a pre-GenAI era and assume that the written artefact provides direct evidence of students' reasoning processes. With GenAI increasingly capable of producing fluent, structurally appropriate and superficially plausible psychological writing, these assumptions require re-evaluation (Cotton et al., 2024; Francis et al., 2025; Tierney et al., 2025).

Williams (2025) identifies two broad strategies that higher education departments are deploying: *mitigation* – reducing opportunities for unacknowledged AI use; and *integration* –actively incorporating GenAI into assessments so that reasoning, critique and metacognitive explanation become the central evidential components. These approaches align with wider work on assessment redesign in digital contexts (Bearman et al., 2023; Bloxham & Boyd, 2007; Boud & Falchikov, 2006), emphasising the need to make assessment processes more transparent and reasoning based.

Mitigation strategies include increasing the use of invigilated assessments, in-class practical tasks, oral examinations or vivas, and scaffolded drafting processes. These formats constrain opportunities for automated text generation and shift emphasis towards students' real-time reasoning, interpretation of results or verbal articulation of methodological decisions. Some departments now require iterative drafts that demonstrate development of ideas, or short reflective commentaries explaining analytic decisions. However, mitigation approaches pose trade-offs: invigilated tests may disadvantage students with anxiety or disabilities; oral components increase staff workload; and procedural checks (e.g., draft logs) can feel bureaucratic to both staff and students (Boud & Falchikov, 2006; Corbin et al., 2025).

Integration strategies, by contrast, attempt to treat GenAI as a legitimate object of disciplinary analysis. Psychology educators have been at the forefront of developing such assessments because they align naturally with the discipline's focus on evidence evaluation, methodological scrutiny and cognitive processes (Chan, 2023; Francis et al., 2025; Lee et al., 2024; Malik et al., 2025). Typical examples include:

- Tasks requiring students to critique AI-generated literature reviews, identifying methodological inaccuracies, conceptual gaps or fabricated citations;
- Assignments where students annotate AI-produced statistical interpretations, explaining errors, misinterpretations or missing assumptions;
- Research-design exercises in which GenAI generates an initial set of hypotheses or interview questions and students must evaluate, refine, or reject them with reference to psychological theory and ethics;
- Reflective components in which students justify when and how they used GenAI, analysing its influence on their reasoning and identifying limitations or cognitive biases.

These integration-oriented assessments foreground evaluation, justification and critical thinking – capacities that GenAI cannot easily automate. They also make students' reasoning processes visible to markers and help develop GenAI literacy. This literacy extends beyond basic tool competence and includes understanding how large language models generate text, recognising common statistical or methodological errors in AI outputs, identifying fabricated data or references, evaluating algorithmic bias, and reflecting on the cognitive risks of over-reliance (Darvishi et al., 2024; O'Donnell et al., 2024).

However, integration strategies also carry practical demands. They require educators to feel confident in using GenAI as a pedagogical tool, to anticipate its typical errors, and to design marking criteria that assess reasoning rather than merely the quality of prose. Staff report that such assessments take longer to mark and require more explicit guidance to students. Moreover, some students initially interpret integration tasks as endorsing unrestricted GenAI use, highlighting the need for clear communication about acceptable practices.

Across the sector, a growing number of psychology departments are piloting mixed-format assessments that blend mitigation and integration, such as combining an invigilated data-analysis task with a take-home critique of an AI-generated research summary. Early evaluations suggest that these formats can maintain assessment validity while building students' evaluative judgement (Francis et al., 2025).

Overall, evolving assessment strategies in psychology reflect a discipline striving to preserve the evidential value of assessments while also recognising the pedagogical opportunities GenAI offers. Scaling these innovations will require institutions to acknowledge psychology's distinctive assessment landscape, provide protected time for redesign and invest in consistent staff development across programmes.

Gaps and Priorities in Staff Development

Staff development is a critical but underdeveloped component of institutional responses to GenAI. Existing training tends to focus on tool demonstrations or high-level policy guidance rather than the disciplinary reasoning, ethical judgement and interpretive skill that psychology educators must exercise when evaluating ambiguous submissions (Hutson et al., 2022). Sector surveys consistently highlight staff anxiety about fairness, clarity and assessment integrity (Lee et al., 2024; Tierney et al., 2025). Colleagues repeatedly ask the same questions: What forms of GenAI use are acceptable in a lab report? How should I respond when reflective writing feels impersonal? How can I mark ethically when authenticity is uncertain?

Effective staff development should therefore integrate ethical reasoning, reflective writing, empirical methods, and GenAI-related scenarios. Case-based learning, using anonymised examples of ambiguous writing, fabricated data, or inconsistent statistical reporting, has proved particularly powerful in early pilots because it engages educators with the tacit judgement required in practice. 'Assessment clinics', shared disclosure templates, and communities of practice can support consistency and provide spaces for collective reflection. From an organisational-learning perspective, such communities function as feedback and knowledge-transfer mechanisms that enable institutions to move beyond individual improvisation towards shared interpretive frameworks and coherent practice (Argyris & Schon, 1978; Senge, 2006). Without these structures, institutions risk perpetuating inconsistency, undermining staff wellbeing and missing opportunities for purposeful adaptation.

Towards a Constructive Vision for Psychology and AI

GenAI presents not only risks but opportunities for psychology educators to deepen student engagement with the epistemic foundations of the discipline (Al-Zahrani & Alasmari, 2024; Chan, 2023). Rather than viewing GenAI solely as a threat, educators can incorporate it meaningfully into teaching in ways that support inclusive, engaging, and reflective learning. Although students frequently use GenAI as a shortcut rather than as a tool for critical engagement (Darvishi et al., 2024), structured activities can help reorient its use. For example, dual-dialogue exercises between students and GenAI can stimulate metacognitive awareness, while classroom tasks that require students to

critique or refine AI-generated drafts draw directly on psychology's strengths in analytical reasoning and reflective practice (Lee et al., 2024; Malik et al., 2025).

Currently, however, most institutional responses remain risk-focused (Bobula, 2024; O'Donnell et al., 2024; Slimi, 2023). Cotton et al. (2024) argue convincingly that institutions should move beyond reactive enforcement towards embedding GenAI literacy, ethical reflection and informed dialogue into curricula, an approach well suited to psychology. These activities directly support the learning outcomes outlined in the APA Guidelines for the Undergraduate Psychology Major (Version 3.0) (American Psychological Association, 2023), particularly those relating to ethical reasoning, scientific inquiry, and the evaluation of evidence – domains that GenAI now makes pedagogically indispensable.

Recommendations and Future Directions

The recommendations that follow draw explicitly on the paper's three conceptual anchors: assessment validity, academic integrity and organisational learning. They aim to support psychology departments in developing coherent and discipline-sensitive responses to GenAI.

Building the Evidence Base: Research Priorities for Psychology and GenAI

To respond effectively to GenAI, psychology requires a stronger empirical foundation that captures the complexity of how AI intersects with disciplinary learning, assessment and identity formation. Existing sector-wide studies tend to homogenise higher education, overlooking how psychology's epistemic values – empirical reasoning, methodological rigour and ethical reflection – shape both opportunities and risks associated with GenAI. A more robust evidence base should therefore pursue several interconnected strands.

First, research should examine *how psychology students actually use GenAI* across different forms of coursework. This includes understanding not only overt uses (e.g., generating summaries or explanations) but also subtle, procedural uses that may be harder to detect, such as drafting, paraphrasing or structuring. Process-tracing or version-history analysis, for instance, could help reveal how GenAI shapes students' cognitive engagement with methodological concepts, ethical reasoning or statistical interpretation.

Second, there is a need to evaluate *the impact of GenAI on learning outcomes, disciplinary thinking and academic identity*. Studies could investigate whether sustained GenAI use erodes conceptual understanding, weakens metacognitive monitoring or alters how students perceive authorship and scientific integrity. Such work would provide important evidence for redesigning assessments and supporting students to develop reflective, ethical engagement with AI.

Third, research should explore *how academic staff are adapting their marking practices, feedback strategies and judgements of authorship in response to GenAI*. Qualitative interviews, marking simulations and cross-institutional comparisons could illuminate how educators navigate uncertainty and how this influences the validity and reliability of marking.

Finally, there is a need to examine *variability across student groups*, including differences by educational background, linguistic diversity and neurodiversity. Understanding these patterns will support equity-sensitive assessment design and help identify whether GenAI amplifies or mitigates existing attainment gaps.

Such research is essential not only for policy development but also for restoring the evidentiary foundations of assessment validity described in the 'Threats to Assessment Validity' section. Building this evidence base will allow psychology departments to make informed, discipline-specific policy decisions, clarify expectations for students, and evaluate whether redesigned

assessments genuinely measure the competencies they target. This research is foundational to the sector's ability to respond coherently to GenAI and to safeguard assessment validity in the long term.

Clarifying Policy and Supporting Fair Practice in Psychology

Current academic integrity policies often lack the precision required for psychology-specific assessments, resulting in inconsistent interpretations and uncertainty for both staff and students. Strengthening policy frameworks is therefore essential to ensuring fairness, protecting academic standards and supporting staff decision-making.

First, institutions should develop *discipline-specific definitions and exemplars* of acceptable and unacceptable GenAI use for each assessment type. For example, what constitutes legitimate support in a lab report versus a reflective assignment? Providing concrete, psychology-focused examples would reduce ambiguity and help students understand expectations.

Second, departments would benefit from *shared templates* for assessment briefs, disclosure statements, authorship declarations and evidence logs. Templates could include guidance on how students should document their use of GenAI and what staff should record when concerns arise. This would promote consistency across modules and reduce the interpretive burden on individual markers.

Third, policies should outline *defensible and proportionate authorship-verification mechanisms*. These may include draft logs, annotated bibliographies showing engagement with sources, short oral verifications, version-history checks, or reflective commentaries explaining analytic choices. Such procedures must protect students from undue suspicion while giving educators confidence in their judgements.

Finally, clearer *escalation pathways* are needed for borderline cases. Policies should specify when concerns should be handled informally, when formal procedures are appropriate, and what level of evidence is required at each stage. This would reduce staff anxiety, support fairness and ensure that decisions are procedurally robust.

Overall, clearer and more discipline-sensitive policy frameworks will help psychology departments navigate the increased ambiguity GenAI introduces into assessment and integrity processes.

Supporting Psychology Educators: Training and Peer Learning

Given the interpretive complexity of evaluating GenAI-influenced work, psychology educators need sustained, practice-oriented professional development. Effective staff support should go beyond tool demonstrations and focus on developing interpretive judgement, disciplinary confidence and shared pedagogical frameworks.

One priority is *case-based learning* using anonymised examples of ambiguous or AI-affected psychology coursework. Reviewing borderline cases collectively allows educators to practise articulating their reasoning, compare interpretations and develop shared standards for identifying problematic patterns (e.g., fabricated data, conceptual inconsistencies and impersonal reflective writing).

Another priority is *training that explicitly integrates GenAI with psychology's core domains*: empirical methods, ethical reasoning, reflective practice and scientific communication. Workshops could explore how GenAI misrepresents statistical results, introduces methodological inaccuracies or produces ethically inconsistent advice, helping staff anticipate the kinds of errors students may reproduce.

Departments should also invest in *communities of practice* that meet regularly to discuss emerging challenges, share strategies, develop exemplars and produce shared guidance. These communities support organisational learning by creating feedback loops through which local innovations can be evaluated, adapted and disseminated across programmes.

Finally, staff development must include *emotional and procedural support*. Many educators experience anxiety about making incorrect decisions, damaging student relationships or misinterpreting policy. Safe spaces for discussion, peer mentoring and clear institutional backing are essential to sustaining morale and ensuring consistent practice.

Reimagining Assessment for a GenAI Era

Rather than focusing solely on preventing GenAI use, psychology has an opportunity to redesign assessments so that they foreground reasoning, judgement, ethical awareness and critical evaluation – competencies that GenAI cannot replicate. This requires both conceptual rethinking and practical experimentation.

One promising direction is the development of *process-oriented assessments* that require students to document their reasoning. For example, students might submit annotated drafts showing how ideas evolved, reflective commentaries explaining methodological decisions or short audio recordings outlining their interpretation of results. These artefacts make thinking visible and help ensure that assessments capture students' own work.

Another direction involves *evaluative tasks* in which students critique, refine or contextualise AI-generated material. Such tasks leverage GenAI's weaknesses – methodological errors, superficial reasoning, fabricated data, as opportunities for students to demonstrate disciplinary competence. This also builds GenAI literacy by helping students recognise limitations and biases in AI outputs.

Psychology departments can also experiment with *hybrid assessments* that blend supervised and unsupervised components, such as an in-person data analysis followed by a take-home theoretical interpretation. These formats limit unacknowledged GenAI use while preserving opportunities for extended reasoning.

Finally, programmes should integrate *GenAI ethics and cognitive science* into research-methods and ethics modules. Topics such as cognitive offloading, algorithmic bias, metacognition and the psychology of trust in automated systems are directly relevant to understanding how students engage with GenAI. Embedding these topics in the curriculum strengthens both ethical judgement and disciplinary understanding.

Taken together, these innovations point towards a future in which assessments are designed not to avoid GenAI but to *withstand* it – by emphasising cognitive processes and disciplinary values that cannot be automated. This approach aligns psychology with contemporary digital scholarship and prepares graduates to use GenAI critically, ethically and creatively.

Conclusion

GenAI is reshaping higher education, and psychology is experiencing these changes with particular intensity. The discipline's reliance on empirical analysis, reflective writing, ethical reasoning and methodological competence means that the challenges discussed throughout this review – unreliable detection, ambiguity in marking and feedback, threats to assessment validity, increased integrity casework, and uneven institutional support – strike at the heart of what psychology assessments are designed to measure. These pressures expose longstanding vulnerabilities in assessment design, highlight the limitations of generic institutional policies, and reveal gaps in staff preparedness.

Yet the expanded recommendations in the fourth section show that psychology is also exceptionally well placed to lead constructive and evidence-informed adaptation. Building a stronger empirical foundation for understanding GenAI use ('Building the Evidence Base: Research Priorities for Psychology and GenAI' section) will allow departments to evaluate whether redesigned assessments genuinely capture disciplinary competencies. Clearer, discipline-specific policy frameworks ('Clarifying Policy and Supporting Fair Practice in Psychology' section) can support fairness, strengthen authorship verification and reduce inconsistency in case handling. Sustained, practice-oriented staff development and communities of practice ('Supporting Psychology Educators: Training and Peer Learning' section) are essential for developing the interpretive judgement and confidence required to navigate GenAI-related ambiguity. Finally, reimagining assessment design to foreground reasoning, evaluative judgement, ethical awareness and process-based evidence ('Reimagining Assessment for a GenAI Era' section) provides a pathway towards assessments that can withstand GenAI rather than be undermined by it.

Taken together, these recommendations point to a coherent strategic response: one that integrates research, policy, training and assessment design rather than treating these components in isolation. Psychology can model this integrated approach because its epistemic values – critical evaluation, methodological rigour, ethical reflection and metacognitive awareness – align closely with the competencies students need to engage responsibly with GenAI. By embedding GenAI literacy into curricula, redesigning assessments to make reasoning visible, and strengthening organisational structures that support staff, psychology departments can create learning environments in which GenAI becomes a tool for deeper engagement rather than a threat to academic integrity.

If institutions invest in these discipline-sensitive strategies, psychology can help shape a future in which GenAI becomes an instrument for cultivating deeper empirical reasoning, stronger ethical judgement, and more resilient academic standards.

ORCID iD

Jean-François Delvenne https://orcid.org/0000-0002-1990-9738

Declaration of conflicting interest

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Abdelaal, N., & Al Sawy, I. (2024). Perceptions, challenges, and prospects: University professors' use of artificial intelligence in education. *Australian Journal of Applied Linguistics*, 7(1), 1309. https://doi.org/10.29140/ajal.v7n1.1309

Acosta-Enriquez, B. G., Arbulú Ballesteros, M. A., Arbulu Perez Vargas, C. G., Orellana Ulloa, M. N., Gutierrez Ulloa, C. R., Pizarro Romero, J. M., Gutierrez Jaramillo, N. D., Cuenca Orellana, H. U., Ayala Anzoategui, D. X., & Lopez Roca, C. (2024). Knowledge, attitudes, and perceived ethics regarding

the use of ChatGPT among generation Z university students. *International Journal for Educational Integrity*, *20*(1), 10. https://doi.org/10.1007/s40979-024-00157-4

Al-Zahrani, A. M., & Alasmari, T. M. (2024). Exploring the impact of artificial intelligence on higher education: The dynamics of ethical, social, and educational implications. *Humanities and Social Sciences Communications*, *11*, 912. https://doi.org/10.1057/s41599-024-03432-4

American Psychological Association. (2023). *APA guidelines for the undergraduate psychology major: Version 3.0*. https://www.apa.org/about/policy/undergraduate-psychology-major.pdf

An, Y., Yu, J. H., & James, S. (2025). Investigating the higher education institutions' guidelines and policies regarding the use of generative AI in teaching, learning, research, and administration. *International Journal of Educational Technology in Higher Education*, *22*, 10. https://doi.org/10.1186/s41239-025-00507-3

Argyris, C., & Schon, D. (1978). *Organizational learning: A theory of action perspective*. Addison-Wesley.

Bearman, M., Nieminen, J. H., & Ajjawi, R. (2023). Designing assessment in a digital world: An organising framework. *Assessment & Evaluation in Higher Education*, *48*(3), 291–304. https://doi.org/10.1080/02602938.2022.2069674

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, *32*, 347–364. https://doi.org/10.1007/BF00138871

Black, R. W., & Tomlinson, B. (2025). University students describe how they adopt AI for writing and research in a general education course. *Scientific Reports*, *15*(1), 8799. https://doi.org/10.1038/s41598-025-92937-2

Bloxham, S., & Boyd, P. (2007). *Developing effective assessment in higher education: A practical guide* (1st ed). Open University Press.

Bobula, M. (2024). Generative artificial intelligence (AI) in higher education: A comprehensive review of challenges, opportunities, and implications. *Journal of Learning Development in Higher Education*, *30*, https://doi.org/10.47408/jldhe.vi30.1137

Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, *31*(4), 399–413. https://doi.org/10.1080/02602930600679050

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, *20*(1), 38. https://doi.org/10.1186/s41239-023-00408-3

Chechitelli, A. (2023, May 23). *AI writing detection update from Turnitin's Chief Product Officer*. Retrieved June 26, 2025, from https://www.turnitin.com/blog/ai-writing-detection-update-from-turnitins-chief-product-officer

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

Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers & Education*, *210*, 104967. https://doi.org/10.1016/j.compedu.2023.104967

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

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hutson, J., Jeevanjee, T., Graaf, V. V., Lively, J., Weber, J., Weir, G., Arnone, K., Carnes, G., Vosevich, K., Plate, D., Leary, M., & Edele, S. (2022). Artificial intelligence and the disruption of higher education: Strategies for integrations across disciplines. *Creative Education*, *13*(12), 3953–3980. https://doi.org/10.4236/ce.2022.1312253

Jisc. (2023; updated 27 November 2025). *Student perceptions of generative AI*. Jisc. Retrieved from Jisc website https://www.jisc.ac.uk/reports/student-perceptions-of-generative-ai

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*, 2522–2549. https://doi.org/10.1111/bjet.13585

Lee, D., Arnold, M., Srivastava, A., Plastow, K., Strelan, P., Ploeckl, F., Lekkas, D., & Palmer, E. (2024). The impact of generative AI on higher education learning and teaching: A study of educators' perspectives. *Computers and Education: Artificial Intelligence*, *6*, 100221. https://doi.org/10.1016/j.caeai.2024.100221

Malik, A., Khan, M. L., Hussain, K., Qadir, J., & Tarhini, A. (2025). AI In higher education: Unveiling academicians' perspectives on teaching, research, and ethics in the age of ChatGPT. *Interactive Learning Environments*, *33*(3), 2390–2406. https://doi.org/10.1080/10494820.2024.2409407

McDonald, N., Johri, A., Ali, A., & Collier, A. H. (2025). Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. *Computers in Human Behavior: Artificial Humans*, *3*, 100121. https://doi.org/10.1016/j.chbah.2025.100121

Newton, P., & Xiromeriti, M. (2024). ChatGPT performance on multiple choice question examinations in higher education. A pragmatic scoping review. *Assessment & Evaluation in Higher Education*, *49*(6), 781–798. https://doi.org/10.1080/02602938.2023.2299059

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. https://doi.org/10.1080/03075070600572090

O'Donnell, F., Porter, M., & Fitzgerald, S. (2024). The role of artificial intelligence in higher education: Higher education students use of AI in academic assignments. *Irish Journal of Technology Enhanced Learning*, *8*(1). https://doi.org/10.22554/szwjfy54

Pallant, J. L., Blijlevens, J., Campbell, A., & Jopp, R. (2025). Mastering knowledge: The impact of generative AI on student learning outcomes. *Studies in Higher Education*, 1–22. https://doi.org/10.1080/03075079.2025.2487570

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). https://doi.org/10.53761/q3azde36

Perkins, M., Roe, J., Vu, B. H., Postma, D., Hickerson, D., McGaughran, J., & Khuat, H. Q., (2024). Simple techniques to bypass GenAI text detectors: Implications for inclusive education. *International Journal of Educational Technology in Higher Education*, *21*, 53. https://doi.org/10.1186/s41239-024-00487-w

QAA. (2024). *Navigating the complexity of the artificial intelligence in higher education*. QAA. Retrieved from QAA website https://www.qaa.ac.uk/docs/qaa/news/quality-compass-navigating-the-complexities-of-the-artificial-intelligence-era-in-higher-education.pdf?sfvrsn=8179b281_11

Russell Group. (2023). Principles on the use of generative AI tools in education. https://www.russellgroup.ac.uk/policy/policy-briefings/principles-use-generative-ai-tools-education

Scarfe, P., Watcham, K., Clarke, A., & Roesch, E. (2024). A real-world test of artificial intelligence infiltration of a university examinations system: A "Turing Test" case study. *PLoS ONE*, *19*(6), e0305354. https://doi.org/10.1371/journal.pone.0305354

Senge, P. (2006). *The fifth discipline: The art and practice of the learning organization*. Random House Books.

Slimi, Z. (2023). The impact of artificial intelligence on higher education: An empirical study. *European Journal of Educational Sciences*, *10*(1), 17–33. https://doi.org/10.19044/ejes.v10no1a17

Tierney, A., Peasey, P. V., & Gould, J. J. M. (2025). Student perceptions on the impact of AI on their teaching and learning experiences in higher education. *Research and Practice in Technology Enhanced Learning*, *20*, 005. https://doi.org/10.58459/rptel.2025.20005

UCISA. (2025, 28 April). *2024 Survey of Digital Education for higher education in the UK*. UCISA. Retrieved from UCISA website https://www.ucisa.ac.uk/news-and-blogs/news/2025/april/de-2024-survey

Williams, A. (2025). Integrating artificial intelligence into higher education assessment. *Intersection: A Journal at the Intersection of Assessment and Learning*, *6*(1), 128–154. https://doi.org/10.61669/001c.131915

Yeadon, W., Inyang, O.-O., Mizouri, A., Peach, A., & Testrow, C. (2023). The death of the short-form physics essay in the coming AI revolution. *Physics Education*, *58*(3), 035027. https://doi.org/10.1088/1361-6552/acc5cf

Author biography

Jean-François Delvenne is an Associate Professor of Cognitive Psychology at the University of Leeds. His research focuses on visual working memory and inter-hemispheric communication, particularly in the context of cognitive ageing. Since 2021, he has served as Academic Integrity Lead within the School of Psychology, where he contributes to assessment policy, academic misconduct processes, and guidance on the pedagogical implications of generative artificial intelligence.