



African Journal of Humanities and Social Sciences

Publisher's Home Page: <https://www.hsspress.com/>



Short Communication

Open Access

Personalized Assessment and Epistemic Justice in AI-Era Humanities Education

Sbusiso Gwala^{1*} 

¹University of Johannesburg, Johannesburg, South Africa. E-mail: gwalasbusiso99@gmail.com

Article Info

Volume 6, Issue 1, February 2026

Received : 25 September 2025

Accepted : 20 January 2026

Published : 25 February 2026

doi: [10.51483/AFJHSS.6.1.2026.1-6](https://doi.org/10.51483/AFJHSS.6.1.2026.1-6)

Abstract

This paper argues that the challenge generative AI poses to humanities education is fundamentally epistemological rather than technological. Drawing on scholarships in assessment theory, epistemic justice, and postcolonial pedagogy, I propose that personalized physical examinations based on students' own submitted work offer a response that is both pedagogically grounded and politically necessary. This approach centers the verification of interpretive engagement rather than text production; privileges situated knowledge over decontextualized outputs and restores assessment as a practice of intellectual accountability. The paper examines the theoretical foundations of this model, its implications for humanities pedagogy in under-resourced contexts, and its contribution to broader conversations about epistemic justice in higher education.

Keywords: *Assessment, Generative AI, Humanities pedagogy, Epistemic justice, Contextualized learning, Academic labor*

© 2026 Sbusiso Gwala. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

1. Introduction: Reframing the Problem

The proliferation of generative AI tools in higher education has generated considerable anxiety, particularly in humanities disciplines where written assignments constitute the primary mode of both learning and assessment. Institutional responses have largely focused on technological solutions, detection software, learning analytics, AI integration strategies, or policy mechanisms such as revised honor codes and assessment redesign toward “AI-proof” tasks (Sullivan *et al.*, 2023; Rudolph *et al.*, 2023).

These responses, however, misapprehend the nature of the challenge. The concern is not primarily that students can generate text without writing it themselves, ghost-writing and academic dishonesty predate AI by centuries. Rather, the fundamental issue is epistemological: how do we verify that students have engaged in the interpretive, analytical, and critical practices that constitute humanities learning when the relationship between text production and intellectual labor has been severed?

* Corresponding author: Sbusiso Gwala, University of Johannesburg, Johannesburg, South Africa. E-mail: gwalasbusiso99@gmail.com

This paper argues that personalized assessment, physical examinations tailored to each student's specific arguments, examples, and analytical frameworks as demonstrated in their coursework, offers a response that addresses this epistemological challenge while advancing broader commitments to epistemic justice in humanities education. Rather than asking whether students produced their own text, we should ask whether they can account for their own thinking. This reorientation has profound implications for how we understand assessment, learning, and knowledge production in humanities disciplines.

2. Theoretical Foundations

2.1. Assessment as Verification of Engagement

Traditional assessment theory distinguishes between formative assessment (feedback for learning) and summative assessment (evaluation of learning outcomes). This binary, however, obscures a third function particularly crucial to humanities disciplines: assessment as verification that students have undertaken the intellectual practices that constitute disciplinary learning (Shay, 2005; Boud, 2000).

In humanities, learning outcomes center on interpretive engagement with texts, construction of evidence-based arguments, critical evaluation of competing frameworks, and contextualized application of theory. These are not knowledge states that can be measured through recall or reproduction, but epistemic practices that must be enacted (Wineburg, 2001; Carter, 2007). The challenge AI poses is not that it can produce acceptable text, but that acceptable text no longer serves as reliable evidence of these underlying practices.

Personalized assessment grounded in students' own coursework directly addresses this verification problem. When examination questions ask students to explain, defend, or extend arguments they themselves advanced in prior assignments, assessment becomes a test of whether they can account for their intellectual choices, a practice that requires having made those choices in the first place. As Rowntree (1987) argues, "If we wish to discover the truth about an educational system, we must look into its assessment procedures". What personalized assessment reveals is whether students engaged in humanistic inquiry or merely produced humanistic text.

2.2. Epistemic Justice and Situated Knowledge

Fricker's (2007) theory of epistemic injustice, particularly testimonial and hermeneutical injustice, provides crucial framing for understanding AI's impact on humanities pedagogy in Global South contexts. When students outsource assignment production to AI tools trained predominantly on Global North academic discourse, they participate in what might be termed "epistemic outsourcing", the displacement of situated, experiential knowledge with decontextualized, algorithmically assembled text.

This displacement is not pedagogically neutral. Generative AI tools exhibit well-documented biases toward Western epistemologies, English-language sources, and Global North examples (Bender et al., 2021; Birhane, 2021). When a South African student writing about transitional justice allows ChatGPT to select case studies, the result typically privileges canonical Global North examples, Nuremberg, Argentina, over local contexts like the Truth and Reconciliation Commission or Marikana. The student's situated knowledge, lived experience, and engagement with local intellectual traditions are erased in favor of algorithmic neutrality that is anything but neutral.

Personalized assessment that holds students accountable for their specific examples, contexts, and applications thus becomes a practice of epistemic justice. It creates pedagogical space where situated knowledge is not only valued but required, where a student cannot pass by reproducing decontextualized Wikipedia summaries but must demonstrate engagement with their own reality. This aligns with calls in postcolonial pedagogy for teaching practices that center rather than marginalize students' epistemological standpoints (Leibowitz et al., 2012; Shahjahan, 2011).

2.3. Metacognitive Accountability

Flavell's (1976) foundational work on metacognition established that learning involves not only acquiring knowledge but developing awareness of one's own cognitive processes. In humanities disciplines, this metacognitive dimension is particularly crucial. Students must not only interpret texts but understand how

and why they arrive at interpretations; not only apply theory but recognize the implications and limitations of their theoretical choices; not only construct arguments but evaluate the strength of their own reasoning.

Traditional assessment often neglects this metacognitive dimension, focusing on product (the essay) rather than process (the thinking that generated it). Personalized examination questions that ask students to explain their intellectual choices “Why did you select this example rather than alternatives?” “What counterarguments did you consider?” “How would your argument change if applied to a different context?”, make metacognitive awareness both visible and assessable.

This has particular importance in the AI context. Students who genuinely engaged in humanistic inquiry can articulate the reasoning behind their choices because they made those choices. Students who outsourced text production to AI cannot provide this account because there was no reasoning process to articulate, only prompt engineering and editorial selection. Metacognitive accountability thus becomes the mechanism through which we distinguish genuine learning from text generation.

3. Pedagogical Implications

3.1. Assessment as Dialogue

Bakhtin's (1981) notion of dialogism that emerges through interaction rather than existing as static property of texts, offers a productive framework for reconceptualizing assessment in the AI era. Traditional examination models treat assessment as monologic, the examiner poses questions, the student provides answers, meaning is determined by correspondence between answer and marking rubric.

Personalized assessment based on students' own work, however, is inherently dialogic. The examination becomes a continuation of an ongoing conversation about the student's developing understanding. The lecturer's question (“You argued X-explain how Y supports this claim”) responds to the student's prior contribution (their assignment argument), which itself responded to course materials and discussions. The student's examination answer extends this dialogue by elaborating, defending, or revising their position. Assessment becomes not the endpoint of learning but a moment within an ongoing process of intellectual exchange.

This dialogic orientation has practical implications for how we design examination questions. Rather than asking students to demonstrate mastery of course content, we ask them to continue thinking with and through that content. Rather than testing recall, we test reasoning. Rather than demanding correct answers, we demand accountable argumentation. The shift from monologic to dialogic assessment realigns evaluation with the actual practices of humanistic inquiry.

3.2. Addressing the Labor Question

Any pedagogical proposal must reckon with the material conditions of its implementation. The personalized assessment model described here is undeniably labor-intensive, requiring tutors to draft individualized questions during the marking process and lecturers to review and verify these questions. In contexts where teaching staff are already stretched thin, where tutors mark 45 assignments on precarious contracts, where permanent academic positions have been casualized, this additional labor cannot be assumed as automatic or unproblematic.

Two responses are necessary. First, we must recognize that this labor is not additional to good humanities teaching, it is constitutive of it. Close engagement with students' work, attention to their specific thinking processes, and individualized feedback are already pedagogical ideals. Personalized assessment doesn't create new obligations so much as formalize and make assessable what excellent teaching already involves. The question is not whether this labor should be done, but whether institutions will resource it adequately.

Second, and more critically, the labor question reveals the political stakes of the AI debate in higher education. If institutions respond to AI by investing in detection software, learning analytics, and technological surveillance rather than in the human labor of teaching, they signal that education is fundamentally about credentialing rather than learning. Conversely, choosing to properly fund personalized, labor-intensive

assessment represents a commitment to pedagogy over policing, to intellectual development over efficiency metrics.

The labor question is thus not a practical obstacle to be solved but a site where competing visions of higher education's purpose become visible. Arguments for personalized assessment must simultaneously be arguments for just compensation of academic labor.

4. Scalability and Equity Considerations

Two significant concerns merit sustained attention: scalability to large courses and equity implications for diverse learners.

Scalability: The personalized assessment model works most naturally in courses of 30-150 students where sustained engagement with individual student work is feasible. In larger courses, modifications are necessary. These might include applying the model only to final examinations rather than mid-term assessments; using stratified sampling where personalized questions are developed for a random selection of students; or creating tiered systems where some questions are personalized while others draw from common pools.

Importantly, partial implementation may still achieve the pedagogical goal. If students know that examination questions may be based on their specific work but cannot predict which students will receive personalized questions, the incentive structure changes. The possibility of accountability for one's arguments may motivate genuine engagement even when such accountability is not universal.

Equity: Personalized assessment privileges certain cognitive and linguistic practices, the ability to articulate one's reasoning process under timed conditions, to engage in abstract metacognitive reflection, to perform intellectual accountability. These practices are not neutral; they reflect academic cultures and may disadvantage students from educational backgrounds where different forms of knowledge demonstrate are valued (Haggis, 2006; Lillis, 2001).

This concern is real but requires careful analysis. The alternative, assessment that does not verify genuine engagement, may worsen equity outcomes by advantaging students who can afford AI subscriptions, edit services, or have stronger English language facility for prompting AI effectively. Personalized assessment levels this playing field by making resource differences less consequential than actual understanding.

Moreover, the concern about timed examination bias can be addressed through accommodations (extended time, alternative formats) without abandoning the core principle, that students should be able to account for their own thinking. The equity question is not whether to implement personalized assessment but how to implement it in ways that recognize diverse learning needs while maintaining standards of intellectual accountability.

5. Epistemic Justice and Contextualized Knowledge

The most compelling argument for personalized assessment may be its alignment with epistemic justice. When AI-generated assignments default to Global North examples, Western theoretical frameworks, and decontextualized analysis, they enact what Santos (2007) terms "epistemicide", the erasure of non-Western knowledge systems and situated ways of knowing.

Consider a concrete example, a student writing about gender-based violence in South Africa might submit an AI-generated essay drawing on American statistics, Canadian policy frameworks, and Nordic interventions. The assignment might receive a passing grade if technical requirements are met. But what has been lost? The student's potential engagement with local feminist scholarship, their understanding of how GBV connects to apartheid's migrant labor system, their analysis of community-based interventions they have witnessed, their critique of how Western frameworks misrecognize cultural context.

Personalized assessment that holds students accountable for their specific examples, their contextual applications, their situated analysis, creates pedagogical pressure toward exactly this kind of engaged, localized knowledge production. It becomes difficult to pass by outsourcing thinking to AI when examination questions require you to explain why your examples matter in your context, how theory illuminates your reality, what your analysis reveals about your society.

This is not parochialism or rejection of global scholarship. Rather, it recognizes that humanities education should prepare students to think with and from their locations, to bring their lived experiences into dialogue with academic discourse, to produce knowledge that matters for their communities. When assessment requires students to account for their intellectual choices, it privileges the situated knowledge that AI cannot generate.

6. Conclusion

The challenge generative AI poses to humanities education is, at its core, a crisis of accountability. How do we maintain standards of intellectual rigor when the relationship between text and thinking has been disrupted? How do we ensure students develop humanistic capabilities when acceptable performance no longer requires humanistic practice?

Personalized assessment based on students' own coursework offers a response grounded in humanities pedagogy's core commitments, that learning involves interpretive engagement, that knowledge is situated and contextual, that arguments require defense and justification. By asking students to account for their own thinking, to explain their examples, defend their frameworks, acknowledge their choices, we create assessment conditions where genuine learning becomes distinguishable from text production.

This approach will not eliminate AI use, nor should it aim to. Students will continue to use these tools for brainstorming, drafting, and editing. What personalized assessment does is ensure that, regardless of which tools supported their work, students must ultimately understand and be able to defend what they produced. The tool becomes servant to thinking rather than substitute for it.

Implementation requires institutional commitment to resource teaching labor adequately, to recognize personalized assessment in workload models, to prioritize pedagogy over surveillance. It requires lecturers and tutors willing to engage deeply with student work and to see examination design as extension of teaching rather than administrative burden. And it requires students willing to accept that accountability for one's arguments is not punitive assessment but the core practice of humanistic inquiry.

If we are serious about humanities education in the AI era, we must stop asking how to prevent students from using these tools and start asking how to ensure they learned to think. Personalized assessment offers one answer, by making thinking visible through the practice of accounting for one's own intellectual choices. In doing so, it restores assessment as a practice of intellectual accountability rather than text verification, which is precisely what humanities education should be.

References

- Bakhtin, M.M. (1981). *The Dialogic Imagination*. University of Texas Press.
- Bender, E.M., Gebru, T., McMillan-Major, A. and Shmitchell, S. (2021). *On the Dangers of Stochastic Parrots: Can Language Models be Too Big?*. *Proceedings of FAccT 2021*.
- Birhane, A. (2021). *Algorithmic Injustice: A Relational Ethics Approach*. *Patterns*, 2(2).
- Boud, D. (2000). *Sustainable Assessment: Rethinking Assessment for the Learning Society*. *Studies in Continuing Education*, 22(2), 151-167.
- Carter, S. (2007). *Ways of Knowing, Doing, and Writing in the Disciplines*. *College Composition and Communication*, 58(3), 385-418.
- Flavell, J.H. (1976). *Metacognitive Aspects of Problem Solving*. *The Nature of Intelligence*, 231-235.
- Fricker, M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press.
- Haggis, T. (2006). *Pedagogies for Diversity: Retaining Critical Challenge Amidst Fears of 'Dumbing Down'*. *Studies in Higher Education*, 31(5), 521-535.
- Leibowitz, B., Bozalek, V., Carolissen, R., Nicholls, L., Rohleder, P. and Swartz, L. (2012). *It's been a Horrible Time but I Think We are Fighting a Just Struggle: Can Service Learning Promote Social Justice?*. *Teaching in Higher Education*, 17(6), 654-663.

- Lillis, T. (2001). *Student Writing: Access, Regulation, Desire*. Routledge.
- Rowntree, D. (1987). *Assessing Students: How Shall We Know Them?*. Kogan Page.
- Rudolph, J., Tan, S. and Tan, S. (2023). ChatGPT: Bullshit Spewer or the End of Traditional Assessments in Higher Education?. *Journal of Applied Learning and Teaching*, 6(1).
- Santos, B.S. (2007). *Beyond Abyssal Thinking: From Global Lines to Ecologies of Knowledges*. *Review*, 30(1), 45-89.
- Shahjahan, R.A. (2011). *Decolonizing the Evidence-Based Education and Policy Movement: Revealing the Colonial Vestiges in Educational Policy, Research and Neoliberal Reform*. *Journal of Education Policy*, 26(2), 181-206.
- Shay, S. (2005). *The Assessment of Complex Tasks: A Double Reading*. *Studies in Higher Education*, 30(6), 663-679.
- Sullivan, M., Kelly, A. and McLaughlan, P. (2023). ChatGPT in Higher Education: Considerations for Academic Integrity and Student Learning. *Journal of Applied Learning and Teaching*, 6(1).
- Wineburg, S. (2001). *Historical Thinking and Other Unnatural Acts*. Temple University Press.