OPEN PRAXIS

Human Agency and Epistemic Authority Under Generative Artificial Intelligence

RESEARCH ARTICLES

MAHMUT ÖZER
MATJAŽ PERC
HANDE TANBERKAN ÖZÇELIK

*Author affiliations can be found in the back matter of this article

INTERNATIONAL COUNCIL FOR OPEN AND DISTANCE EDUCATION

ABSTRACT

This study examines the epistemic, pedagogical, and institutional ruptures introduced by generative artificial intelligence in educational assessment and evaluation, as well as in scholarly publishing. As large language models (LLMs) achieve increasingly high levels of fluency and coherence, it becomes harder to determine who the relevant agent is behind learning outcomes and academic texts. In response, education systems and peer-reviewed publishing have shown a growing tendency to delegate assessment, evaluation, and oversight to AI-based tools. We argue that this shift is not merely a technical adjustment but a structural transformation that erodes human responsibility and epistemic authority by assigning both production and judgment to closely related algorithmic systems. By discussing the limitations of AI detection tools—especially their false-positive risks—the article highlights the ethical and epistemic problems that arise when academic integrity is reduced to the formal features of text. In educational contexts, LLM-supported assignments and examinations can obscure students' cognitive effort and weaken the connection between learning and achievement. In scholarly publishing, the same dynamic encourages a surveillance-oriented posture that treats style as evidence of authorship while failing to secure reliability, originality, or conceptual contribution. Against this backdrop, the study argues that the solution lies not in more sophisticated surveillance, but in redesigning assessment and evaluation to re-anchor production and judgment in human oversight. For education, we propose a framework that structurally separates the learning phase from the evaluation phase: LLMs may be used as complementary tools during learning, while evidence of learning is produced under controlled conditions and human supervision. For scholarly publishing, we advocate replacing detection regimes with transparent disclosure of AI use, while keeping final evaluation grounded in human peer review. We conclude that this approach can protect pedagogical validity and academic integrity while restoring human responsibility and epistemic standing in the age of generative artificial intelligence.

CORRESPONDING AUTHOR:
Mahmut Özer
Turkish Grand National Assembly, Türkiye
mahmutozer2002@yahoo.com

KEYWORDS:
large language models; artificial intelligence; epistemology; education; research; assessment; evaluation; epistemic authority; generative artificial intelligence; accountability

TO CITE THIS ARTICLE:
Özer, M., Perc, M., & Özçelik, H. T. (2026). Human Agency and Epistemic Authority Under Generative Artificial Intelligence. *Open Praxis*, *18*(3), pp. 495–505. DOI: https://doi.org/10.55982/openpraxis.18.3.1149

INTRODUCTION

Generative artificial intelligence applications such as ChatGPT have become embedded in many domains of everyday life, including education and healthcare (Bozkurt, 2023; İlikhan et al., 2024; 2025; Özer, 2024a; Özer and Perc, 2024; Özer et al., 2024a; 2024b). Large language models (LLMs) can generate text from prompts, summarize and revise documents, extract key points, and translate across languages at high speed. For students, teachers, academics, journalists, and other professional groups, these capabilities have rapidly become part of routine practice, positioning LLMs as widely used supportive tools (Bozkurt, 2024; Özer, 2024a; 2024b; Pavlik, 2023). As LLM performance improves, adoption continues to grow; yet with this growth, the relationship between users and these tools has also begun to resemble a form of dependency that often remains unnoticed (Perc & Özer, 2025).

From the perspective of education systems, this development creates three interrelated problems. First, student use is shifting from assistance to substitution: LLMs are increasingly used in place of the student rather than by the student (Özer, 2024a). Second, this substitution can weaken cognitive capacities that education aims to cultivate especially critical thinking and memory by reducing sustained engagement with difficult tasks (Ahmad et al., 2023; Gerlich, 2025; Kosmyna et al., 2025; Özer et al., 2025; Perc& Özer, 2025; Stadler et al., 2024). Third, education systems still lack scalable assessment and evaluation frameworks that can reliably determine the true agent behind student work produced with these tools (Suna& Özer, 2025; 2026; Tanberkan et al., 2024). In short, the agent of production is becoming ambiguous (Floridi, 2023), and agency in the human–machine relationship is increasingly displaced toward the machine.

This ambiguity has prompted new strategies in both educational assessment and scholarly communication. In education, institutions are experimenting with methods intended to re-establish student agency under conditions where conventional take-home work no longer provides trustworthy evidence of learning. Notably, classical oral examinations and dialogue-based assessments are re-emerging as ways to test students' understanding and to attribute work to an accountable agent (Barba& Stegner, 2026; Furze et al., 2024; Roe et al., 2025).

A closely related tension is emerging in scholarly publishing. LLMs can produce manuscripts with persuasive fluency, structural coherence, and apparent alignment with the literature. Yet such surface quality does not guarantee accuracy, originality, or conceptual contribution. Peer review functions precisely here: it provides human testimony about a manuscript's reliability and about whether it constitutes a meaningful scientific contribution. As LLM use expands, however, there is growing pressure to automate parts of this gatekeeping process. If peer review is reduced to AI-based screening and recommendation, human testimony is displaced, and epistemically fragile texts may circulate more easily (Thorp, 2023). The risk is not simply automation of production, but automation of oversight in ways that reproduce the same limitations as the tools used to produce the text.

In response, AI detection tools have been developed to infer whether an academic text contains "hidden authorship" by LLMs. Beyond their technical shortcomings, these tools introduce deeper epistemic and ethical complications (Giray, 2024; Rashidi et al., 2023). They presume that human writing and LLM output are statistically distinguishable and therefore rely on surface-level probabilistic signals such as predictability, word distributions, repetition patterns, sentence-length variance, and syntactic regularity (Giray, 2024; Kusal et al., 2021). A common method uses perplexity as a proxy for textual "surprise," assuming that AI-generated text tends to be less surprising than human-authored prose (Elek et al., 2025; Xu & Sheng, 2024). Some approaches also anticipate the use of embedded "watermarks" that are detectable by specific algorithms (Amrit & Singh, 2022).

The first tension is that these tools can end up policing a writing style rather than authorship. Text that is regular, normative, and highly aligned with disciplinary conventions may be flagged as suspicious, meaning that human writing is evaluated through the machine's statistical averages (Giray, 2024). This is especially problematic in academia, where writing is already standardized by disciplinary norms, editorial practices, and reviewer feedback. As standardization increases, the risk that detection tools will misattribute authorship increases as well. Consequently, fully human-authored academic prose can take forms that resemble LLM output.

This helps explain why well-structured, norm-compliant texts—particularly those produced by authors with strong command of disciplinary conventions—may be flagged. Given the generally low performance of these detection tools (Giray, 2024; Gotoman et al., 2025; Kar et al., 2025; Legaspi et al., 2024), concerns intensify when false positives disproportionately affect non-native English–speaking scholars (Fraser et al., 2025). Subramaniam (2023), for example, tested OpenAI's detection tool across languages and reported low accuracy for non-English texts such as Arabic and Hindi, indicating linguistic bias. More recent work similarly reports weak performance for texts written in Persian (Shamsi et al., 2026).

A further concern is that some tools identify "AI-like" features in texts written before LLMs were widely available, suggesting that what is being detected may be stylistic regularity rather than AI use. One early study analyzed 14,400 genuine abstracts from five leading journals published between 1980 and 2023. While most were correctly identified as human-authored, approximately 5–10% were incorrectly labeled as AI-generated with high confidence (sometimes exceeding 90%) (Rashidi et al., 2023). The issue is not whether such error rates are mathematically tolerable (Giray, 2024), but the asymmetry of harms: even a single false positive can trigger accusations that damage careers and reputations despite genuine authorship.

In education, the analogous problem is that LLM-assisted assignments, essays, projects, and even exam-like outputs may no longer reliably indicate what students know, how they reason, or how they connect concepts (Barba & Stegner, 2026; Furze et al., 2024; Perkins et al., 2024; Roe et al., 2025). If assessment focuses primarily on polished products, it risks mistaking algorithmic fluency for student understanding. In this sense, assessment and peer review play comparable roles: both are intended to serve as external tests of production. When evaluation itself is delegated to AI-based systems, this externality is weakened.

These developments converge into a new paradox: the same class of systems increasingly supports text production and then is tasked with judging that production. Because the systems that generate and the systems that evaluate share similar epistemic constraints, evaluation can become an echo of production rather than an independent check (O'Neil, 2016; Quattrociocchi et al., 2026; Özer, 2024c; Özer, 2025a; 2025b; Özer & Perc, 2025). Historically, the purpose of assessment in education and peer review in science has been precisely to test outputs from an external standpoint. The present shift risks eliminating that standpoint.

Accordingly, this study examines generative AI in education and scholarly publishing through the lens of human agency—how humans can remain responsible producers and responsible judges—and proposes a framework designed to preserve assessment validity and academic integrity without collapsing oversight into automated surveillance.

RE-CENTERING HUMAN AGENCY IN EDUCATION ASSESSMENT

Breaking out of the current vicious cycle requires more than a technical fix; it demands an epistemic, institutional, and moral repositioning. The solution is not to marginally refine algorithms, but to re-anchor both production and judgment in human agency. Here, Thorp (2023), Editor-in-Chief of Science, captures the challenge succinctly:

> "Many concerns relate to how ChatGPT will change education. It certainly can write essays about a range of topics. I gave it both an exam and a final project that I had assigned students in a class I taught on science denial at George Washington University. It did well finding factual answers, but the scholarly writing still has a long way to go. If anything, the implications for education may push academics to rethink their courses in innovative ways and give assignments that aren't easily solved by AI. That could be for the best."

Thorp's emphasis on "assignments that aren't easily solved by AI" is a direct challenge to the prevailing form of exams and coursework. It calls for assessment designs that foreground interpretation, justification, and defensible reasoning—forms of evaluation that make students' thinking visible rather than merely inspecting the surface quality of a final text. At its core, this is a demand to preserve human beings as the primary agents of learning and to re-conceptualize assessment accordingly.

If evaluation concentrates on the finished product alone, LLMs will predictably dominate that space. Yet learning (like scientific inquiry) is constituted by process: effort, conceptual change, error correction, and the development of reasons. Assessment must therefore be redesigned to privilege evidence of thinking over polished output. What matters is not only what a student submits, but why they chose a particular approach, how they evaluated alternatives, and what justifications they can provide. This implies shifting assessment from a one-off judgment to a practice that can include structured reflection, iterative feedback, and, where appropriate, brief defenses or interviews that test understanding rather than stylistic fluency.

Several responses in the recent literature illustrate both the promise and the limits of this direction. Roe et al. (2025) propose "assessment twins" as a way to protect validity in higher education: the same learning outcome is assessed through two complementary forms of evidence, such as pairing a written assignment with an in-class task, oral defense, applied problem, peer interaction, or guided discussion. The aim is cross-validation—reducing the chance that LLM-supported production can masquerade as learning—while strengthening content and construct validity. However, the authors also acknowledge practical constraints, including scalability in large classes and the inequities that can arise for students who experience anxiety in oral formats. It is also unclear how students with weaker oral skills, who do not rely on these tools, would fare in such a model. Similarly, Barba and Stegner (2026) propose a conversation-based examination and show its implementation in a cohort of 58 software engineering students using a question bank, decision trees, pre-determined hints, and standardized rubrics.

A broader and more flexible proposal is the Artificial Intelligence Assessment Scale (AIAS), a five-level model designed to structure permissible AI use in education (Perkins et al., 2024). A pilot implementation demonstrates how the AIAS levels—from "no AI" to "full AI"—can be operationalized in practice (Furze et al., 2024). At Level 1, assessment follows a classical logic: the goal is to measure what students know and can do independently, prioritizing secure conditions and individual performance. At Level 2, students may use LLMs, but they may not include AI-generated content directly in the final submission; AI is confined to supportive functions, keeping the locus of intellectual production with the student. From Level 3 onward, the object of assessment shifts. At Level 3, students are evaluated less on authorship of a final text and more on their capacity to revise and develop their thinking coherently with AI support; the focus moves from writing performance to reasoning and editing, accompanied by disclosure of AI contributions. At Level 4, assessment targets judgment: students complete tasks using these tools but are evaluated on their ability to critique AI output for accuracy, bias, and limitations. At Level 5, AI use becomes the default, and assessment focuses on students' capacity to collaborate with AI—selecting tools, directing them effectively, and achieving defined objectives.

From Level 3 onward, the AIAS framework shifts attention away from simply patching the vulnerabilities introduced by LLMs and toward integrating AI use across contexts in a structured way. As its proponents argue, the approach aims to build contextual awareness and provide practical guidance for ethically sensitive integration. A key limitation, however, concerns educational equity: not all students have equal access to these tools, and unequal access can translate into unequal performance opportunities.

These developments also expose a deeper paradox in the current moment. LLMs were widely expected to reduce educators' workload, enrich learning environments, and expand personalized learning at scale. Yet many emerging assessment responses function primarily as compensatory mechanisms—designed to repair validity threatened by substitutive uses of LLMs. In practice, systems often attempt to remedy technology-induced fragilities by returning to labor-intensive, pre-digital methods (e.g., oral verification). This is not necessarily misguided, but it signals that the dominant problem is not the existence of LLMs, but the way their capabilities make substitution attractive.

Although the immediate issue may appear to be "misuse," the affordances of LLMs—speed, fluent production, and perceived performance advantages—actively incentivize substitutive use. This incentive is reinforced by structural pressures within education systems: high-stakes assessment, efficiency demands, and scalability constraints. For that reason, it is unrealistic to rely primarily on individual ethical appeals or expectations of "proper use."

Unless assessment architectures are redesigned so that complementary use is enabled while substitution is structurally disincentivized, dependency will predictably persist. As models continue to improve—becoming faster, smoother, and more persuasive—this tendency will likely strengthen rather than fade.

At the same time, the alternatives are more problematic. Comprehensive bans are rarely enforceable, and treating LLM-produced outputs as straightforward evidence of learning would undermine assessment validity. Complementary use therefore remains the only sustainable middle path, but it requires deliberate pedagogical and institutional design and a willingness to bear its costs. The central question is less whether such an architecture is possible, and more whether institutions are willing to abandon the easiest solutions in order to preserve educational integrity.

A second problem follows from the first: the burden placed on educators. If safeguarding validity depends on oral verification or "double" assessment, educator workload increases, and the system risks drifting back toward a teacher-centered model. This is a genuine tension. Student-centered learning presupposes students as active agents, yet substitutive LLM use undermines that agency by allowing evidence of learning to be outsourced to machines. Meanwhile, some integrity-preserving solutions can unintentionally intensify the educator's role as the primary verifier and gatekeeper.

For this reason, we propose a framework that strengthens student agency while remaining scalable. The core design is to separate the learning phase from the evaluation phase. Assignments in which LLM use is permitted should be clearly communicated in advance; students should be allowed to use all resources—including LLMs—during preparation. However, the production of final evidence of learning should occur under supervised conditions within the school environment and under teacher oversight. In this model, LLMs can support learning, but they cannot substitute for the student at the point where competence is evidenced.

This approach offers several advantages. First, it preserves the benefits associated with LLMs during learning—personalization, speed, and exposure to diverse explanations while relocating evaluation to an equitable setting that protects validity. Second, it reduces the trust-eroding suspicion that students' work is primarily machine-produced: seeking help during preparation is not criminalized, and expectations are clear. Third, it avoids the anxiety and inequities that oral examinations can generate for some students. Fourth, it limits the need for redundant verification: teachers are not required to reassess work they have already evaluated through additional oral reconfirmation, which helps contain workload.

Most importantly, the framework directly mitigates the three risks identified at the outset. Students who rely on substitution during preparation will be unable to demonstrate competence under supervised conditions and will face the consequences of that choice. Because substitution becomes an unreliable strategy, the drift toward diminished critical thinking and memory is also constrained. Students' role as the primary agents of learning is thereby reinforced rather than rhetorically asserted.

Implementing this architecture requires strengthening educators as designers and judges of learning. Their fluency with these tools must exceed that of students if they are to structure valid assessment in an AI-saturated environment. This educational challenge has a direct parallel in scholarly publishing: the question is not whether AI appears in academic work, but how institutions preserve responsibility, transparency, and human judgment when it does. The next section addresses this parallel in the context of scientific authorship and peer review.

RESTORING HUMAN JUDGEMENT IN SCHOLARLY PUBLISHING

As discussed above, current AI detection tools do not reliably identify generative AI involvement in scientific writing. In practice, they tend to flag a particular style—text that is highly regular, unsurprising, and norm-compliant—because they retrospectively search for statistical resemblance to the kinds of outputs LLMs commonly produce. This is a crucial distinction: these tools do not perform causal inference about how a text was produced; they estimate formal similarity. Yet similarity is not evidence of use. A well-edited manuscript, a text shaped by disciplinary conventions, or writing refined through legitimate editorial processes can easily resemble the statistical profile of contemporary LLM output. In such a regime, academic integrity

is displaced from an ethical and scholarly judgment to a matter of crossing a machine-defined threshold, amplifying false-positive risks—especially for scholars whose native language is not English.

This produces an epistemic distortion. Surveillance-based detection reduces authorship to surface features while ignoring what authorship actually entails: intention, the selection and interpretation of sources, conceptual contribution, argumentative responsibility, and accountability for error. Academic integrity is not primarily about how a text looks; it concerns how it is produced, what intellectual work it embodies, and what forms of scrutiny it has undergone. For this reason, a growing body of scholarship argues that detection-centered regimes are scientifically fragile and ethically difficult to defend. When academic prose is already standardized and LLMs can imitate that standardization, decisions that rest on stylistic resemblance are structurally unreliable.

The problem deepens when viewed through the lens of epistemic authority. In modern scholarship, legitimacy is established through institutional processes grounded in human judgment: peer review, editorial deliberation, reasoned justification, and openness to critique. A detection regime relocates that authority to an opaque algorithm that offers limited transparency and limited avenues for contestation. Over time, this shift risks normalizing a new regime in which knowledge becomes legitimate not through deliberation and critique, but through algorithmic approval based on similarity metrics.

As these tools spread, they can also reshape academic behavior. If "AI-like" is operationalized as fluent, regular, and convention-following, then disciplined writing—often the result of expertise and editorial refinement—becomes a liability. Authors may respond defensively by writing against the detector rather than for clarity and rigor: deliberately roughening prose, disrupting fluency, or making strategic stylistic choices to avoid suspicion. In the long run, this pressures scholarly communication away from epistemic justification and toward performative maneuvering, degrading the norms that academic integrity is supposed to protect.

A further institutional risk is that algorithmic authority can assert itself retroactively. When detection systems classify texts written before the widespread availability of LLMs as "machine-generated," they reveal the underlying mechanism: not evidence-based attribution, but reclassification by formal criteria. This undermines trust not only in the tools but also in the wider evaluative environment, fostering a culture in which human production is treated as presumptively suspicious and must be "validated" by machines.

For these reasons, the stakes are not limited to false positives or technical error rates. What is at stake is the gradual positioning of AI as both producer and judge, with corresponding erosion of human responsibility and accountability. The appropriate response is therefore not to build more sophisticated detectors, but to redesign scholarly integrity frameworks so that responsibility, transparency, and human judgment remain central.

Accordingly, we propose shifting the focal point from the text's surface features to the production process. Instead of treating integrity as a binary contest between confession and detection ("did you use AI or not?"), it should become a structured disclosure practice ("how, where, and to what extent was AI used?"). Within such a framework, the use of LLMs is not treated as a hidden violation to be policed, but as a methodological choice that must be openly declared—specifying the stage of use (e.g., outlining, translation, editing), the purpose, and the boundaries. This moves oversight away from probabilistic suspicion and toward accountable transparency, while protecting authors from unjust accusations rooted in stylistic resemblance.

Crucially, transparency does not weaken evaluation; it clarifies what peer review should be evaluating. Questions of understanding, conceptual novelty, methodological soundness, and the ability to defend claims under critique cannot be outsourced to detectors. They require human judgment. For this reason, peer review and editorial assessment should be reinforced rather than displaced: the final evaluation of scholarly contribution must remain grounded in reasoned human scrutiny, not in automated similarity scoring. In short, the sustainable path is not a surveillance regime that attempts to infer authorship from textual fingerprints, but a human-centered publishing process that makes AI use visible, keeps responsibility with the author, and re-anchors legitimacy in human review and justification.

DISCUSSION

Generative artificial intelligence is now widely used across educational levels—from primary education to higher education—and throughout academic work by students, educators, and scholars. The central problem this creates for assessment and evaluation is the question of agency and ownership: who is responsible for the produced work, and what exactly does the work evidence? In response to this uncertainty, institutions increasingly turn to assessment mechanisms intended to identify authorship and integrity. Yet the growing use of AI-based tools to perform that very evaluation introduces a further and more structural risk.

When LLMs are used to produce texts in education and scholarly publishing, and when screening, similarity checks, and even evaluative recommendations are likewise performed by algorithms, the human agent is progressively displaced. The machine comes to occupy both roles: producer and evaluator. This is not simply a debate about whether AI "should" be used in assessment and evaluation. The more fundamental question concerns the delegation of judgment: to whom, and within what limits, are we transferring decision-making authority? When production and evaluation are concentrated within the same technological rationality, epistemic authority shifts away from humans and toward machines—implicating responsibility, accountability, and moral agency.

A key consequence of this shift is that evaluation becomes increasingly oriented toward the final product while the process that should justify the product becomes opaque. LLM-supported outputs can be polished and coherent, but that surface quality does not reveal whether a student formed conceptual connections, confronted misunderstanding, or developed justificatory competence. Over time, this risks eroding cognitive effort, self-regulation, and critical thinking. The same logic applies in scholarly contexts: texts may look methodologically and rhetorically "correct" while still being unreliable, derivative, or conceptually thin. If evaluation is reduced to the detection of stylistic signals or to automated endorsements, integrity becomes a matter of appearance rather than accountable production.

For these reasons, this study argues that preventing generative AI from becoming a decision-making authority is the core normative task in both educational assessment and scholarly publishing. AI can support learning and writing, but final judgment, justification, and meaning-making must remain human responsibilities. Otherwise, science and education risk converging on a regime of performance that appears efficient and objective while becoming increasingly hollow.

To break this cycle, we propose an approach grounded in clarifying the locus of responsibility. The agent of the produced text, the rendered judgment, and the attributed success must remain human. AI may be used as support, but moral and institutional responsibility for error, misjudgment, and success cannot be delegated. Without this clarity, neither the seriousness of learning nor the reliability of scientific production can be sustained.

In education, the proposed framework rests on a structural separation between learning and assessment. During learning and preparation, the use of AI tools can be permitted as complementary aids. However, the production of evidence of learning should occur under controlled conditions and human oversight. This design aims to make student agency visible again—by requiring demonstration, explanation, and justification in contexts where substitution is constrained—while avoiding the costs and scalability limits of continuous oral verification or other labor-intensive monitoring regimes. The goal is not surveillance, but validity: ensuring that what is assessed genuinely corresponds to learning.

A recent study evaluating the implications of generative artificial intelligence through the lens of the Community of Inquiry (CoI) framework—originally proposed as an important model for understanding and designing online and blended learning experiences (Bozkurt, 2019; Garrison et al., 2000, 2001)—shows that two indispensable conditions for preserving the integrity of this framework are safeguarding human agency and ensuring that the learning process remains fundamentally dialogical (Stenbom et al., 2026). In this way, the relationship between learners and technological tools such as AI is shifted toward a pedagogical ground that is dialogic and open to development. In other words, students are prevented from becoming passive, one-way recipients of technologies that are highly prone to being positioned as epistemic authorities. As learners develop the capacity to evaluate, critique, and refine the outputs provided by these

tools, the CoI framework can be preserved without requiring any fundamental transformation, while its presences are enhanced through the incorporation of AI tools.

Our proposal is consistent with this perspective.

Similarly, Terry Anderson, who developed the Interaction Equivalency Theorem as a flexible model for creating meaningful learning pathways (Anderson, 2003, 2008), argues that among the three fundamental forms of interaction for learners—content, peers, and teachers—artificial intelligence technologies will affect content most profoundly and peer interaction only partially, while emphasizing that good teachers should continue to inspire and motivate their students and, to ensure this, proposing as a necessity the revival of the tradition of oral examinations for assessing learners even in distance education, and indeed recommending that we seek ways to implement such practices on a broader scale (Anderson & Bozkurt, 2025).

Therefore, these tools can be actively employed prior to assessment, but their use should proceed not in a one-directional manner, but rather through processes of evaluation, critique, and iterative improvement. During assessment itself, however, students do not have access to these tools. Consequently, the fundamental responsibility of educators is not to allow technological tools to evolve unchecked toward the destinations they would naturally reach if left to themselves, but rather to position them where they ought to be in ways that preserve the spirit of the CoI framework and to ensure the continuity of this orientation.

In scholarly publishing, we propose replacing surveillance-oriented detection regimes with disclosure-based transparency. Because attributing authorship through a text's formal features is both technically fragile and ethically risky, integrity should be evaluated less by how a text appears and more by how it was produced. Within this model, AI use is treated as a methodological choice that must be openly declared, specifying at what stage, for what purpose, and within which boundaries it was employed. Final evaluation and accountability remain grounded in human peer review and reasoned justification, thereby preventing epistemic authority from shifting toward algorithmic approval and restoring responsibility to the human agent.

In addressing this issue, the developers of the Technological Pedagogical Content Knowledge (TPACK) framework—an influential model for understanding and designing educational processes and learning environments (Koehler & Mishra, 2005, 2009; Mishra & Koehler, 2006)—have recently incorporated the notion of "pride in work" as a pedagogical strategy within the framework (Mishra & Bozkurt, 2026). Individuals who take pride in their work not only accept responsibility for what they do, but also endeavor to maximize the effort they devote to improving the quality of their work. Our proposal is consistent with and reinforces this perspective. As this sense of pride is strengthened, individuals are more likely to use the opportunities offered by these tools in a responsible and transparent manner. In turn, their sense of ownership and their capacity for agency are continually reinforced and sustained.

Taken together, these proposals re-center assessment and evaluation on human judgment while allowing AI to remain a legitimate tool for learning and drafting. By separating learning from evaluation in education and replacing detection with disclosure in publishing, the framework aims to safeguard pedagogical validity and academic integrity while re-securing human responsibility and epistemic standing in the age of generative artificial intelligence.

DATA ACCESSIBILITY STATEMENT

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

SUSTAINABLE DEVELOPMENT GOALS (SDGs)

This study is linked to the following SDG(s): Quality education (SDG 4) and Reduced inequalities (SDG 10).

ETHICS AND CONSENT

Because this paper is conceptual in tis natüre, ethical review was waived in this study.

FUNDING INFORMATION

M.P. was supported by the Slovenian Research and Innovation Agency (Grant No. P1-0403).

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS (CRediT)

Mahmut Özer: Conceptualization, investigation, writing—original draft preparation, writing—review and editing; Matjaz Perc: supervision, writing—original draft preparation, writing—review and editing; Hande Tanberkan Özçelik: investigation, writing— original draft preparation. All authors have read and agreed to the published version of the manuscript.

AUTHOR AFFILIATIONS

Mahmut Özer orcid.org/0000-0001-8722-8670
Turkish Grand National Assembly, Türkiye

Matjaž Perc orcid.org/0000-0002-3087-541X
University of Maribor, Slovenia; Community Healthcare Center Dr. Adolf Drolc Maribor, Slovenia; Korea University, Republic of Korea; Kyung Hee University, Republic of Korea

Hande Tanberkan Özçelik orcid.org/0000-0001-7142-5397
Baskent University, Türkiye

REFERENCES

Ahmad, S. F., Han, H., Alam, M. M., Rehmat, M. K., Irshad, M., Arraño-Muñoz, M., & **Ariza-Montes, A.** (2023). Impact of artificial intelligence on human loss in decision making, laziness and safety in education. *Humanities and Social Sciences Communications*, *10*(311). https://doi.org//10.1057/s41599-023-01787-8

Amrit, P., & **Singh, A. K.** (2022). Survey on Watermarking Methods in the Artificial Intelligence domain and beyond. *Computer Communications*, *188*, 52–65. https://doi.org/10.1016/j.comcom.2022.02.023

Anderson, T. (2003). Getting the mix right again: An updated and theoretical rationale for interaction. *The International Review of Research in Open and Distributed Learning*, 4(2). https://doi.org/10.19173/irrodl.v4i2.149

Anderson, T. (2008). Towards a theory of online learning. In T. Anderson (Ed.), *Theory and Practice of Online Learning* (pp. 109–119). AU Press. https://doi.org/10.15215/aupress/9781897425084.004

Anderson, T., & **Bozkurt, A.** (2025). A critically informed conversation with Terry Anderson: Visions on the next generation of online and distance education. *Open Praxis*, *17*(4), 830–835. https://doi.org/10.55982/openpraxis.17.4.976

Barba, L. A., & **Stegner, L.** (2026). The conversational exam: a scalable assessment design for the ai era. *arXiv*:2601.10691. https://doi.org/10.48550/arXiv.2601.10691

Bozkurt, A. (2019). Intellectual roots of distance education: a progressive knowledge domain analysis. *Distance Education*, *40*(4), 497–514. https://doi.org/10.1080/01587919.2019.1681894

Bozkurt, A. (2023). Unleashing the Potential of Generative AI, Conversational Agents and Chatbots in Educational Praxis: A Systematic Review and Bibliometric Analysis of GenAI in Education. *Open Praxis*, *15*(4), 261–270. https://doi.org/10.55982/openpraxis15.4.609

Bozkurt, A. (2024). Why Generative AI Literacy, Why Now and Why it Matters in the Educational Landscape? Kings, Queens and GenAI Dragons. *Open Praxis*, *16*(3), 283–290. https://doi.org/10.55982/openpraxis.16.3.739

Elek, A., Yildiz, H. S., Akca, B., Oren, N. C., & **Gundogdu, B.** (2025). Evaluating the efficacy of perplexity scores in distinguishing ai-generated and human-written abstracts. *Academic Radiology*, *32*(4), 1785–1790. https://doi.org/10.1016/j.acra.2025.01.017

Floridi, L. (2023). AI as agency without intelligence: on chatgpt, large language models, and other generative models. *Philosophy & Technology*, *36*, 15. https://doi.org/10.1007/s13347-023-00621-y

Fraser, K. C., Dawkins, H., & **Kiritchenko, S.** (2025). Detecting AI-generated text: Factors influencing detectability with current methods. *Journal of Artificial Intelligence Research*, *82*, 2233–2278. https://doi.org/10.1613/jair.1.16665

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*, *40*(4). https://doi.org/10.14742/ajet.9434

Garrison, D. R., Anderson, T., & **Archer, W.** (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, *2*(2), 87–105. https://doi.org/10.1016/S1096-7516(00)00016-6

Garrison, D. R., Anderson, T., & **Archer, W.** (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of Distance Education*, *15*(1), 7–23. https://doi.org/10.1080/08923640109527071

Gerlich, M. (2025). AI Tools in Society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, *15*(1), 6. https://doi.org/10.3390/soc15010006

Giray, L. (2024). The Problem with false positives: AI detection unfairly accuses scholars of AI plagiarism. *The Serials Librarian*, *85*. https://doi.org/10.1080/0361526X.2024.2433256

Gotoman, J. E. J., Luna, H. L. T., Sangria, J. C. S., Santiago, C. S., Jr., & **Barbuco, D. D.** (2025). Accuracy and reliability of AI-generated text detection tools: A literature review. *American Journal of IR 4.0 and Beyond*, *4*(1), 1–9. https://doi.org/10.54536/ajirb.v4i1.3795

İlikhan, S., Özer, M., Perc, M., Tanberkan, H., & **Ayhan, Y.** (2025). Complementary use of artificial intelligence in healthcare. *Medical Journal of Western Black Sea*, *9*(1), 7–17. https://doi.org/10.29058/mjwbs.1620035

İlikhan, S., Özer, M., Tanberkan, H., & **Bozkurt, V.** (2024). How to mitigate the risks of deployment of artificial intelligence in medicine? *Turkish Journal of Medical Science*, *54*(3), 483–492. https://doi.org/10.55730/1300-0144.5814

Kar, S. K., Bansal, T., Modi, S., & **Singh, A.** (2025). How sensitive are the free AI-detector tools in detecting AI-generated texts? A comparison of popular AI-detector tools. *Indian journal of Psychological Medicine*, *47*(3), 275–278. https://doi.org/10.1177/02537176241247934

Koehler, M. J., & **Mishra, P.** (2005). What happens when teachers design educational technology? The development of technological pedagogical content knowledge. *Journal of Educational Computing Research*, *32*(2), 131–152. https://doi.org/10.2190/0ew7-01wb-bkhl-qdyv

Koehler, M. J., & **Mishra, P.** (2009). What is technological pedagogical content knowledge? *Contemporary Issues in Technology and Teacher Education*, *9*(1), 60–70. https://www.learntechlib.org/d/29544/

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X., Beresnitzky, A. V., Braounstein, I., & **Maes, P.** (2025). *Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for an essay writing task*. arXiv:2506.08872. https://doi.org/10.48550/arXiv.2506.08872

Kusal, S., Patil, S., Kotecha, K., Aluvalu, R., & **Varadarajan, V.** (2021). AI-Based emotion detection for textual big data: techniques and contribution. *Big Data and Cognitive Computing*, *5*(3), 43. https://doi.org/10.3390/bdcc5030043

Legaspi, J. B., Licuben, R. J. O., Legaspi, E. A., & **Tolentino, J. A.** (2024). Comparing AI detectors: evaluating performance and efficiency. *International Journal of Science and Research Archive*, *12*(2), 833–838. https://doi.org/10.30574/ijsra.2024.12.2.1276

Mishra, P., & **Bozkurt, A.** (2026). TPACK in the age of context and artificial intelligence: Punya Mishra on the "Wicked Problem" of AI and why algorithms can't replace context. *Open Praxis*, *18*(2), 398–407. https://doi.org/10.55982/openpraxis.18.2.1093

Mishra, P., & **Koehler, M. J.** (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record: The Voice of Scholarship in Education*, *108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increase Inequality and Threatens Democracy*. Crown Books.

Quattrociocchi, W., Capraro, V., & **Perc, M.** (2026). *Epistemological Fault Lines Between Human and Artificial Intelligence*. arXiv: 2512.19466. https://doi.org/10.31234/osf.io/c5gh8_v1

Özer, M. (2024a). Potential benefits and risks of artificial intelligence in education. *Bartın University Journal of Faculty of Education*, *13*(2), 232–244. https://doi.org/10.14686/buefad.1416087

Özer, M. (2024b). Impact of ChatCPT on scientific writing. *İnsan ve Toplum Dergisi*, *14*(3), 210–217. https://doi.org/10.12658/E0002

Özer, M. (2024c). Is artificial intelligence hallucinating? *Turkish Journal of Psychiatry*, *35*(4), 333–335. https://doi.org/10.5080/u27587

Özer, M. (2025a). The epistemic crisis of artificial intelligence in the context of Nafs al-Amr. *Journal of Economy Culture and Society*, *72*, 254–264. https://doi.org/10.26650/JECS2025-1780354

Özer, M. (2025b). Can mathematical models be weapons of mass destruction? *Reflektif Journal of Social Sciences*, *6*(1), 259–268. https://doi.org/10.47613/reflektif.2025.212

Özer, M., & **Perc, M.** (2024). Human complementation must aid automation to mitigate the unemployment effects due to AI technologies in the labor market. *Reflektif Journal of Social Sciences*, *5*(2), 503–514. https://doi.org/10.47613/reflektif.2024.176

Özer, M., & **Perc, M.** (2025). Creative alineation in art due to artificial intelligence. *Reflektif Journal of Social Sciences*, *6*(3), 1105–1118. https://doi.org/10.47613/reflektif.2025.258

Özer, M., Perc, M., & **Suna, H. E.** (2024a). Artificial intelligence bias and the amplification of inequalities in the labor market. *Journal of Economy, Culture and Society*, *69*, 159–168. https://doi.org/10.26650/JECS2023-1415085

Özer, M., Perc, M., & **Suna, H. E.** (2024b). Participatory management can help AI ethics adhere to the social consensus. *Istanbul University Journal of Sociology*, *44*(1), 221–238. https://doi.org/10.26650/SJ.2024.44.1.0001

Özer, M., Tanberkan, H., & **Perc, M.** (2025). Artificial intelligence threatens critical thinking in education systems. *Journal of Higher Education and Science*, *15*(2), 157–164. https://doi.org/10.5961/highereduSci.1747885

Pavlik, J. V. (2023). Collaborating with ChatGPT: considering the implications of generative artificial intelligence for journalism and media education. *Journalism & Mass Communication Educator*, *78*(1), 84–93. https://doi.org/10.1177/10776958221149577

Perc, M., & **Özer, M.** (2025). Disappearing minds in the age of artificial intelligence. *İnsan ve Toplum Dergisi*, *15*(3), 1–9. https://doi.org/10.12658/E0004

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

Rashidi, H. H., Fennell, B. D., Albahra, S., Hu, B., & **Gorbett, T.** (2023). The ChatGPT conundrum: Human-generated scientific manuscripts misidentified as AI creations by AI text detection tool. *Journal of Pathology Informatics*, *14*, 100342. https://doi.org/10.1016/j.jpi.2023.100342

Roe, J., Perkins, M., & **Giray, L.** (2025). *Assessment Twins: A protocol For AI-Vulnerable summative assessment*. arXiv:2510.02929. https://doi.org/10.48550/arXiv.2510.02929

Shamsi, A., Wang, T., Amraei, M., & **Raju, N. V.** (2026). Evaluating AI text detection tools for distinguishing human-written from AI-generated abstracts in Persian language journals of library and information science. *Acta Informatica Pragensia*, *15*(1), 126–134. https://doi.org/10.18267/j.aip.293

Stadler, M., Bannert, M., & **Sailer, M.** (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, *160*. https://doi.org/10.1016/j.chb.2024.108386

Stenbom, S., Garrison, D. R., & **Bozkurt, A.** (2026). Augmenting inquiry, preserving the core: Stenbom and Garrison on AI's role and human-centered learning within the Community of Inquiry (CoI) framework. *Open Praxis*, *18*(1), 181–191. https://doi.org/10.55982/openpraxis.18.1.1042

Subramaniam, R. (2023). Identifying text classification failures in multilingual AI-Generated content. *International Journal of Artificial Intelligence & Applications*, *14*(5), 57–63. https://doi.org/10.5121/ijaia.2023.14505

Suna, H., & **Özer, M.** (2026). Assessment in higher education in the era of generative AI: A mediator for new opportunities and unique challenges. In *Reimagining Curriculum and Assessment in the Age of Generative AI*, p.241–273. IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3373-7579-3.ch008

Suna, H. E., & **Özer, M.** (2025). The human complimentary usage of AI and ML for fair and unbiased educational assessments. *Chinese/English Journal of Educational Measurement and Evaluation*, *6*(1), 1–21. https://doi.org/10.59863/YPKL4338

Tanberkan, H., Özer, M., & **Gelbal, S.** (2024). Impact of artificial intelligence on assessment and evaluation approaches in education. *International Journal of Educational Studies and Policy*, *5*(2), 139–152. https://doi.org/10.5281/zenodo.14016103

Thorp, H. H. (2023). ChatGPT is fun, but not an author. *Science*, *379*(6630), 313. https://doi.org/10.1126/science.adg7879

Xu, Z., & **Sheng, V. S.** (2024). Detecting AI-Generated Code Assignments Using Perplexity of Large Language Models. *Proceedings of the AAAI Conference on Artificial Intelligence*, *38*(21), 23155–23162. https://doi.org/10.1609/aaai.v38i21.30361

TO CITE THIS ARTICLE:
Özer, M., Perc, M., & Özçelik, H. T. (2026). Human Agency and Epistemic Authority Under Generative Artificial Intelligence. *Open Praxis*, *18*(3), pp. 495–505. DOI: https://doi.org/10.55982/openpraxis.18.3.1149

Submitted: 30 January 2026
Accepted: 10 June 2026
Published: 04 August 2026