Epistemic authority and generative AI in learning spaces: rethinking knowledge in the algorithmic age

Binny Jose[1]*, Anu Cleetus[2], Bindhu Joseph[2], Lumy Joseph[3], Benymol Jose[3] and Amruth K. John[3]

[1]Department of Health and Wellness, Marian College Kuttikkanam Autonomous, Kuttikkanam, India, [2]St Joseph College of Teacher Education for Women, Kochi, India, [3]Department of Computer Applications, Marian College Kuttikkanam Autonomous, Kuttikkanam, India

Introduction: generative AI and the crisis of epistemic authority

Generative AI's emergence in learning environments has triggered a potent transformation in how we produce, consume, and verify knowledge. With tools like ChatGPT and other large-scale language models, classrooms are now populated with technologies that assist students in producing essays, solving math problems, summarizing texts, and constructing arguments. These systems bring new efficiencies—but also raise deeper pedagogical questions. What happens when students begin treating machine-generated outputs as epistemic authorities? How might this reshape the traditional roles of teachers and learners in determining what counts as knowledge?

Historically, the classroom has been a space not only for the transfer of information but for co-constructing knowledge through discussion, inquiry, and critique (Freire, 1970; Mejía-Arauz and Wells, 2005). The teacher's authority has rested not just in content expertise but in guiding learners through processes of meaning-making grounded in human judgment, experience, and ethical inference (Van Oers and Dobber, 2015). However, the rise of generative AI—capable of producing fluent, instantaneous, and confident responses—is subtly reconfiguring where authority in learning resides, often without institutional or pedagogical mediation.

This article investigates the epistemological consequences of this shift. We examine how generative AI challenges epistemic agency, disrupts traditional justifications for knowledge, and reshapes the teacher-student dynamic. We explore the idea of AI as a surrogate knower, analyze the hidden curriculum embedded in AI systems, and propose strategies for reclaiming human-centered authority and agency in the classroom.

While much of the current discourse focuses on academic dishonesty or automation of labor, we argue for a more foundational conversation: one that interrogates the restructuring of knowledge itself. Drawing from empirical studies and philosophical models, we contend that generative AI is not merely a new tool, but a new actor in the epistemic space—one whose authority must be understood, challenged, and situated within a pedagogy that resists passive consumption and fosters critical discernment.

AI as surrogate knower: the collapse of justification?

The kind of epistemic presence introduced by generative AI systems such as ChatGPT is a new type—one that appears fluent, confident, and perpetually available (Anjum et al., 2023). To many students, such instruments are not merely aids, but "superior knowers" (Rastogi and Lawati, 2024). This signals a basic epistemological shift: from the collaborative construction of knowledge to a framework that privileges immediate, polished output (Schei et al., 2024). The implications of this are profound.

More recent studies suggest that students increasingly see ChatGPT as something considerably more substantial than support software, using it as the central, trustworthy source of learning (Rastogi and Lawati, 2024). Such trust tends to skirt critical reflection on biases or the model's limits. This is the automation bias, one in which AI results come to be trusted too much, even when users question or recognize errors (Benz et al., 2022; Nguyen, 2023). Such tendencies compromise ground-level epistemic practices—such as evidence assessment, source triangulation, and epistemic modesty.

Above all, this transformation is not merely practical, but philosophical. According to Coeckelbergh (2025) AI systems impact not only beliefs, but belief revision itself—the reconfiguring of mechanisms in which people adopt, reject, or modify claims to knowledge. In the classroom, this implies students will update their knowledge in response to algorithmic authority, without recourse to further justification or reflection. This failure of normative epistemic checks further obfuscates the distinction between tool and epistemic agent.

Here, AI is no longer a neutral assistant. It becomes a substitute for a knower—reshaping what students regard as justified knowledge and who they consider experts. The educator's challenge is not to reject these systems, but to reassert epistemic agency in classrooms now cohabited by fluent but non-sentient interlocutors.

Erosion of teacher authority and the reconfiguration of pedagogy

The increasing dominance of AI-created information in schoolrooms threatens not just what is taught, but even who can teach. As generative AI becomes a quiet dialogue partner in learning environments, learners start adjusting their perceptions of whose information carries more weight: that of their educator or that of the algorithm.

Rising research suggests that learners increasingly seek out ChatGPT to support or contradict teacher feedback, an indication that AI is being used more and more as an epistemological counterpoint to human teaching (Gordon and Foucault, 1980). For example, learners report being more willing to seek feedback from ChatGPT than from peers or teachers, even though they are not convinced it is reliable (Marquart and Bruhn, 2025). The research shows that most learners prefer feedback from AI because it is immediate, clear, and fluent, and prefer it to traditional teacher feedback. Specifically, feedback delivered through AI has been associated with reduced writing anxiety and improved fluency, suggesting that speed and polish contribute to its assumed

authority (Wang, 2024). The dynamic can contribute to a gradual undermining of pedagogical trust.

This shift is compounded when instructors themselves begin using AI tools to assess student work or generate feedback—whether to save time or adhere to institutional demands for efficiency. When students recognize that the same generative system is evaluating their performance and assisting them in completing assignments, the boundary between "student" and "teacher" knowledge production becomes blurred. If both are deferring to the same tool, epistemic authority becomes further displaced from the human educator to the algorithm. This not only disrupts traditional authority but creates confusion around who ultimately evaluates understanding.

Additionally, the teacher's authority—long based in their ability to facilitate understanding and provide ethical judgment—is now more often subjected to algorithmic benchmarking. Students may easily assume that the AI is "objective" and that the teacher is "biased," particularly in situations requiring interpretation or critique. These comparisons are not often done consciously, though they shape classroom affective and intellectual dynamics.

Such epistemic displacement is especially precarious in institutional contexts where teachers—particularly women and racialized faculty—have historically been denied full authority in the eyes of students. Research in higher education repeatedly demonstrates that women and faculty of color are more likely to be challenged or discredited by students, especially in contexts involving authority, grading, or political content (Chávez and Mitchell, 2020). In such settings, the rise of generative AI as an "objective" voice risks reinforcing, rather than counteracting, these credibility gaps. The result is a dual undermining: first by existing structural biases, and second by the technological displacement of authority.

This change in some contexts has resulted in a kind of role reversal, wherein teachers are asked to justify their reasoning in contrast to machine output—frequently without any institutional guidance or support. This not only burdens teachers with defending their professional relevance but can also result in a retreat from rich pedagogical practices toward those that conform more closely to algorithmically scorable content.

The outcome is a quiet remaking of authority: from educator expertise to machine-made credibility. Absent a clear pedagogical counter-narrative, this shift threatens to hollow out the teacher's epistemological role.

The algorithmic hidden curriculum

All learning tools carry embedded pedagogies—messages about what is considered knowledge, how to learn, and which voices are seen as authoritative. Generative AI models, though often portrayed as value-neutral or purely functional, are no exception. Their output contains assumptions, perpetuates dominant discourses, and conditions user expectations in subtle but enduring ways.

Recent research indicates that AI-generated content can reinforce cultural and linguistic biases, which students might unknowingly reproduce in their academic work. For instance, AI applications have been shown to replicate gender bias in classroom content, influencing students' writing and framing

decisions (Mattiazzi, 2025). Similarly, language generation patterns affect how students engage with topics and formulate arguments (Singh, 2024). These tools do not merely offer information—they teach rhetorical norms, foreground specific topics, and promote particular moral frameworks—implicitly shaping what students come to see as academically appropriate.

Consider a student tasked with writing a position paper on climate justice. Turning to ChatGPT for help, the student receives a highly structured, Western liberal humanist framing of the issue—centered on policy reform and individual behavior. Without recognizing this as a partial epistemic stance, the student may internalize this framework as the correct or default way to approach such topics. In this way, AI systems implicitly frame not only what is "true," but also what kinds of reasoning and perspectives are permissible.

Scholars have cautioned that AI systems tend to replicate Western epistemologies embedded in their training data, thereby marginalizing plural or Indigenous ways of knowing. As Lewis et al. (2024) explain, most existing AI models reflect rationalist frameworks that systematically exclude non-Western knowledge systems. Similarly, Ofosu-Asare (2024) identifies the persistence of cognitive imperialism in AI, whereby Eurocentric reasoning is privileged unless explicitly counterbalanced through design. Thus, the AI functions as an epistemic filter—valorizing certain forms of knowledge while silencing others.

The hidden curriculum is not merely a question of content, but also of cognitive stance. Generative AI tends to favor fluency, assertiveness, and certainty over ambiguity or discomfort. Its quick, confident responses often discourage deeper questioning, modeling a style of knowledge delivery that rewards speed over critical reflection (Li, 2024; Orlanda-Ventayen, 2024). Gradually, this encourages learners to shift from inquiry-based learning toward reproduction of polished, preformed answers.

Unless explicitly addressed, this epistemic conditioning risks undermining teachers' efforts to cultivate ambiguity tolerance, intellectual humility, and reflective skepticism—qualities that are essential to pluralistic, democratic learning environments.

Reclaiming epistemic agency: toward a pluralistic and reflective pedagogy

In light of algorithmic epistemology's encroachment on traditional learning structures, it is more necessary than ever to cultivate intentional, reflective, and persistent learners. Revitalizing epistemic agency includes not just empowering students to use AI tools effectively, but to interrogate, contextualize, and critique them. It is about reaffirming human agency in the production of knowledge—not as passive recipients of algorithmic information, but as active, situated interpreters.

Automation bias tends to lead users to over-rely on AI systems, especially when those systems present information with fluency and confidence. Research shows that assertive, polished output encourages uncritical acceptance—even in cases of clear error or contradiction (Horowitz and Kahn, 2023; Kutza et al., 2024). This misplaced trust discourages learners from asking critical questions or engaging in epistemic self-reflection.

What is needed in response is the deliberate infusion of epistemic vigilance in educational practices. Specifically, students can be directed to compare AI responses against peer or instructor responses, identifying gaps, assumptions, as well as rhetorical differences. "Trust audits," in a controlled manner, can encourage students to query when and why they feel most ready to trust the AI. Journals or reflection essays can invite learners to note the degree to which their thoughts varied after being presented machine-generated material. These interventions move well beyond digital literacy—they aim to recuperate a more dialogic, evaluative relationship toward knowledge.

These practices illustrate what it is to be an "epistemic mentor"—an instructor who doesn't impart information, but educates for discernment. An epistemic mentor shows students how to proceed in uncertainty, estimate credibility, and comprehend that knowledge is disputed and provisional. This encompasses assisting learners in recognizing their own positionality as well as the sociotechnical circumstances that determine the instrumentations at their command. Instead of protecting students from the impact of AI, educators can support critical engagement with it— questioning not only what is transmitted as knowledge, but why it is set out in this manner, and by whom.

These capabilities ground epistemic agency, defined here as the power of the learner to question, warrant, and claim to know responsibly. In practice, this means verifying AI-generated content, cross-checking sources, and the reliance on human judgement, particularly when matters of interpretation, ethics, or context arise. Educators need institutional support not merely to deliver material well, but to be recognized as epistemic agents in their own right—offerors of rich, sophisticated thought in the digitally rich learning environment.

Conclusion: resisting epistemic automation

With AI integrated into daily pedagogical practices, the teacher is confronted with the pressing epistemological question: What is teaching—or knowledge—when machines issue automatic, confident, and fluent answers at will? This article has been looking at the way generative AI reassigns classroom power, reconfigures students' notion of expertise, and reifies an unobtrusive curriculum that values fluency at the expense of depth and economy at the expense of inquiry. But the question at issue here is not merely technological, it is at base philosophical and relational.

Artificial intelligence programs are revolutionizing the manner in which students come to know, what they hold true, and whose voices they listen to as authoritative. If not checked, these advancements risk emboldening the pattern of passive epistemic consumption, negating the student's agency as well as the instructor's role as mentor to critical, reflective thought.

In response, this paper has made three core contributions. First, we analyzed how AI tools function as surrogate knowers, subtly collapsing justification norms and privileging algorithmic output over dialogic reasoning. Second, we highlighted how both student and institutional behavior can erode the teacher's epistemic authority, particularly in contexts already shaped by structural

inequities. Third, we outlined strategies for reclaiming epistemic agency—through reflective pedagogy, classroom practices that foster critical AI literacy, and a renewed model of the teacher as epistemic mentor.

For educators, researchers, and institutions engaging with digital technologies in education, the challenge is not to reject AI but to reframe its place in the learning process. This includes equipping students to interrogate AI-generated claims, fostering awareness of cognitive bias, and designing learning environments that support epistemic plurality.

The future of education will not be defined solely by what AI can generate, but by what human learners—guided by reflective educators—choose to question, interpret, and reimagine. Reclaiming epistemic agency is not only a pedagogical imperative but a democratic one. In the algorithmic age, learning how to think critically is inseparable from learning how to resist automation as the default mode of knowledge.

Author contributions

BinnJ: Conceptualization, Writing – original draft, Writing – review & editing. AC: Writing – original draft, Writing – review & editing. BindJ: Writing – original draft, Writing – review & editing. LJ: Writing – original draft, Writing – review & editing. BeJ: Writing – original draft, Writing – review & editing. AJ: Writing – original draft, Writing – review & editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Anjum, G., Choubey, J., Kushwaha, S., and Patkar, V. (2023). AI in education: evaluating the efficacy and fairness of automated grading systems. *Int. J. Innov. Res. Sci. Eng. Technol.* 12, 9043–9050. doi: 10.15680/IJIRSET.2023.1206161

Benz, C., Kühl, N., Satzger, G., and Schemmer, M. (2022). On the influence of explainable AI on automation bias. *ArXiv*. doi: 10.48550/arXiv.2204.08859

Chávez, K., and Mitchell, K. M. (2020). Exploring bias in student evaluations: gender, race, and ethnicity. *Polit. Sci. Polit.* 53, 270–274. doi: 10.1017/S1049096519001744

Coeckelbergh, M. (2025). AI and epistemic agency: how ai influences belief revision and its normative implications. *Soc. Epistemol.* 1–13. doi: 10.1080/02691728.2025.2466164

Freire, P. (1970). *Pedagogy of the Oppressed*. New York, NY: Herder and Herder. Available online at: https://scispace.com/search (Accessed June 16, 2025).

Gordon, C., and Foucault, M. (1980). *Selected Interviews and Other Writings* (NewYork, NY), 1972–1977.

Horowitz, M., and Kahn, L. (2023). Bending the automation bias curve: a study of human and AI-based decision making in national security Contexts. *ArXiv* 68. doi: 10.1093/isq/sqae020

Kutza, J.-O., Dissemond, J., Busch, D., Babitsch, B., Przysucha, M., Erfurt-Berge, C., et al. (2024). Automation bias in AI-decision support: results from an empirical study. *Stud. Health Technol. Inform.* 317, 298–304. doi: 10.3233/SHTI240871

Lewis, J. E., Whaanga, H., and Yolgörmez, C. (2024). Abundant intelligences: placing AI within Indigenous knowledge frameworks. *AI Soc.* 40, 2141–2157. doi: 10.1007/s00146-024-02099-4

Li, W. (2024). "The influence of artificial intelligence on the education system," in *Lecture Notes in Education Psychology and Public Media* (Littlemore). doi: 10.54254/2753-7048/65/20240158

Marquart, F., and Bruhn, T. (2025). *Student Experiences of ChatGPT as a Feedback Tool in Higher Education*. Copenhagen: Tidsskriftet Læring Og Medier (LOM).

Mattiazzi, M. (2025). AI and language: new forms for old discriminations? A case study in google translate and canva. *Feminismo/s* 45, 118–138. doi: 10.14198/fem.2025.45.05

Mejía-Arauz, R., and Wells, G. (2005). Toward dialogue in the classroom: learning and Teaching through Inquiry. *Papeles de Trabajo Sobre Cultura, Educación y Desarrollo Humano* 1.

Nguyen, T. (2023). ChatGPT in medical education: a precursor for automation bias? *JMIR Med. Educ.* 10. doi: 10.2196/preprints.50174

Ofosu-Asare, Y. (2024). Cognitive imperialism in artificial intelligence: counteracting bias with indigenous epistemologies. *AI Soc.* 40, 1–17. doi: 10.1007/s00146-024-02065-0

Orlanda-Ventayen, C. (2024). "Empowering education through transformative role of artificial intelligence (AI) in teaching and learning: educators' perspective and research trends," in *2024 9th International Conference on Information Technology and Digital Applications (ICITDA)* (Sembilan), 1–5.

Rastogi, A., and Lawati, A. G. A. A. H. A. (2024). "Understanding the acceptance of ChatGPT by HEI's students for knowledge enhancement," in *2024 1st International Conference on Innovative Engineering Sciences and Technological Research (ICIESTR)* (Muscat), 1–8. doi: 10.1109/ICIESTR60916.2024.10798141

Schei, O. M., Møgelvang, A., and Ludvigsen, K. (2024). Perceptions and use of AI chatbots among students in higher education: a scoping review of empirical studies. *Educ. Sci.* 14:922. doi: 10.3390/educsci14080922

Singh, A. (2024). "Diverse yet biased: towards mitigating biases in generative AI (student Abstract)," in *AAAI'24/IAAI'24/EAAI'24: Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence and Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence and Fourteenth Symposium on Educational Advances in Artificial Intelligence*, 23653–23654. doi: 10.1609/aaai.v38i21.30512

Van Oers, B., and Dobber, M. (2015). The role of the teacher in promoting dialogue and polylogue during inquiry activities in primary education. *Mind Cult. Activity* 22, 326–341. doi: 10.1080/10749039.2014.992545

Wang, D. (2024). Teacher- versus AI-generated (poe application) corrective feedback and language learners' writing anxiety, complexity, fluency, and accuracy. *Int. Rev. Res. Open Distrib. Learn.* 25, 37–56. doi: 10.19173/irrodl.v25i3.7646