Nova Science Press

Journal of Psychology & Education

Vol. 1, No. 6 (2026)

Generative Artificial Intelligence and Student Agency: Reclaiming Questioning, Knowledge Construction, and Responsibility in Education

Hongxuan Ran[1]*

1 Soochow University, Suzhou 215006, Jiangsu, China

*Corresponding author: Hongxuan Ran; email: rhx1202@163.com

Journal of Psychology & Education • Vol. 1, No. 6 (2026)

DOI: https://doi.org/10.66581/hs48qf06

Received 26 April 2026 • Accepted 18 July 2026 • Published 31 July 2026

CITATION

Ran, H. (2026). Generative Artificial Intelligence and Student Agency: Reclaiming Questioning, Knowledge Construction, and Responsibility in Education. Journal of Psychology & Education, 1(6), 86-114. https://doi.org/10.66581/hs48qf06

Abstract

Generative artificial intelligence (GenAI) has rapidly entered educational practice as a tool for explanation, drafting, feedback, assessment support, and personalized learning assistance. Much of the early debate has focused on whether GenAI improves efficiency or threatens academic integrity. This article argues that the deeper educational question is whether GenAI use preserves or erodes student agency. From an educational perspective, agency is not reducible to student choice, motivation, or technical competence; it refers to learners' capacity to pose meaningful questions, construct knowledge through disciplined inquiry, regulate their own learning, exercise judgment, and assume responsibility for their intellectual work. Drawing on educational theory, self-regulated learning, knowledge-building pedagogy, cognitive offloading, and recent research on AI-assisted learning, this article develops a conceptual account of how GenAI may both support and weaken student agency. It identifies three core risks: the displacement of question-formation by prompt optimization, the substitution of fluent AI output for knowledge construction, and the diffusion of responsibility in authorship and judgment. The article then proposes an agency-preserving pedagogy of GenAI use built around six principles: problem-posing before prompting, process transparency, critical verification, scaffolded rather than substitutive assistance, assessment redesign, and teacher-guided ethical reflection. The central claim is that GenAI should not be treated as either a threat to be banned or an efficiency tool to be adopted uncritically. Its educational value depends on whether it is embedded in learning environments that strengthen students' epistemic, metacognitive, dialogic, and responsible agency.

Keywords: generative artificial intelligence, student agency, student subjectivity, knowledge construction, AI literacy, educational technology

Introduction

Generative artificial intelligence has changed the conditions under which students read, write, search, summarize, translate, solve problems, and receive feedback. Unlike earlier digital tools that primarily stored, transmitted, or organized information, GenAI systems can produce coherent text, simulate dialogue, generate examples, revise drafts, and respond adaptively to user prompts. For educators, this creates a genuine instructional opportunity. Students can receive immediate explanations, compare alternative answers, practice argumentation, and overcome some barriers associated with blank-page anxiety or unequal access to tutoring. Recent education guidance therefore treats AI literacy and responsible use as emerging competencies rather than marginal technical skills (UNESCO, 2023, 2024a, 2024b).

Yet educational technologies rarely enter schools and universities as neutral instruments. They reorganize attention, redefine effort, reshape assessment, and alter the relation between learners, teachers, knowledge, and institutions. The question is therefore not simply whether GenAI can improve learning efficiency. It clearly can, under some conditions. The more fundamental question is what kind of learner is being formed when students learn with systems that can answer before they have fully questioned, compose before they have fully reasoned, and polish before they have fully struggled with meaning.

This article argues that the central educational issue raised by GenAI is student agency. The term agency is often used loosely to mean autonomy, freedom of choice, or active participation. In a stronger educational sense, however, student agency involves the capacity to define problems, pursue inquiry, regulate learning strategies, evaluate evidence, revise understanding, and take responsibility for judgments and actions (Bandura, 2001; Emirbayer & Mische, 1998; Zimmerman, 2002). Education is not merely the transfer of correct answers. It is the formation of persons capable of asking better questions, participating in shared knowledge practices, and acting with judgment in uncertain situations (Biesta, 2015, 2017; Dewey, 1938).

The current debate often moves between two inadequate positions. One position treats GenAI as a productivity accelerator and assumes that faster writing, quicker feedback, and personalized explanations are educational goods in themselves. The other position treats GenAI mainly as a threat to academic integrity and responds by banning or policing its use. Both positions capture part of the truth but miss the deeper pedagogical problem. The value of GenAI depends less on the tool itself than on the educational relations and tasks within which it is used. A student who uses GenAI to interrogate assumptions, test counterarguments, and revise a self-generated explanation may become more agentic. A student who uses the same tool to bypass question formation, outsource reasoning, and submit polished but unexamined text may become less agentic.

The purpose of this theoretical article is to develop a framework for understanding and preserving student agency in GenAI-mediated education. The article first clarifies the meaning of student agency from an educational perspective. It then analyzes the affordances of GenAI as a learning support and the mechanisms through which agency may be weakened. Finally, it proposes a pedagogy of agency-preserving GenAI use, with implications for curriculum, assessment, teacher practice, and institutional norms. The argument is deliberately educational: the concern is not any external agenda, but the preservation of the learner as a questioning, meaning-making, responsible subject.

From Tool Adoption to Educational Subjectivity

Educational debates about technology often begin with adoption. Institutions ask which tools should be allowed, which platforms should be purchased, and which policies should regulate use. These questions are necessary but insufficient. An educational perspective begins with a different question: What forms of learning and personhood does a technology make easier, and what forms does it make harder? This question moves beyond technological determinism. GenAI does not automatically improve or damage education. It becomes educationally significant through the practices, norms, assessments, and relationships into which it is inserted (Fawns, 2022; Selwyn, 2019).

Student subjectivity refers to the learner's position as a person who does not merely receive instruction but participates in the formation of meaning. In classical educational theory, the student is not a passive container of information. Dewey (1916, 1938) argued that learning emerges through experience, inquiry, and reflective reconstruction. Vygotsky (1978) emphasized that cognitive development occurs through mediated activity and social interaction. Freire (1970) criticized educational models that reduce learners to recipients of deposited knowledge. Although these traditions differ, they share a basic premise: education becomes genuine when learners participate in meaning-making rather than merely consuming answers.

GenAI challenges this premise because it can produce the appearance of meaning-making without requiring the learner to undertake the corresponding intellectual work. A generated essay may display coherence without personal understanding. A generated explanation may sound plausible without being critically examined. A generated research question may be grammatically elegant but disconnected from the learner's own puzzlement. The danger is not simply cheating. The deeper danger is a split between performance and formation: students may learn to produce artifacts that resemble learning while the capacities that education is supposed to cultivate remain underdeveloped.

This distinction matters because educational success should not be defined solely by output quality. If a tool enables students to submit cleaner prose but reduces their ability to generate questions, tolerate uncertainty, evaluate evidence, and revise their thinking, then the tool has improved performance while weakening education. Conversely, if GenAI is used to make reasoning visible, stimulate comparison, invite critique, and support revision, it can become a scaffold for deeper learning. The same technology can either externalize thinking in a productive way or replace thinking in a harmful way. The pedagogical issue is therefore not whether GenAI is present, but how agency is distributed between student, teacher, technology, and task.

Conceptualizing Student Agency in GenAI-Mediated Learning

Student agency is a multidimensional educational construct. It includes motivation and choice, but it cannot be reduced to either. A student may choose to use a tool frequently and still be unagentic if the tool determines the direction, substance, and evaluation of the work. Conversely, a student may work within clear instructional constraints and still be highly agentic if those constraints require independent judgment, self-regulation, and responsibility. Agency is not the absence of guidance; it is the capacity to act meaningfully within mediated conditions (Bandura, 2001; Emirbayer & Mische, 1998).

Four dimensions of agency are especially relevant to GenAI-mediated education. The first is epistemic agency: the capacity to ask questions, identify gaps in understanding, seek evidence, distinguish warranted claims from plausible assertions, and contribute to knowledge construction (Bereiter, 2002; Scardamalia & Bereiter, 2006). The second is metacognitive agency: the capacity to monitor one's own understanding, select strategies, evaluate progress, and adjust effort (Winne & Hadwin, 1998; Zimmerman, 2002). The third is dialogic agency: the capacity to participate in conversation with teachers, peers, texts, and tools in ways that extend rather than close inquiry. The fourth is responsible agency: the capacity to own one's intellectual decisions, disclose assistance appropriately, and remain accountable for claims made in one's work.

These dimensions clarify why GenAI use can be educationally ambiguous. A student who asks GenAI to produce an answer may show technical skill in prompt writing but little epistemic agency if the student has not formulated the problem or evaluated the response. A student who asks GenAI to challenge a draft argument, identify missing evidence, or simulate opposing viewpoints may strengthen agency because the tool becomes part of a reflective inquiry cycle. The difference lies not in the presence of AI assistance but in whether the student remains the author of the question, the judge of the evidence, and the responsible owner of the final claim.

This account also avoids a romantic view of unaided learning. Human learning has always been mediated by tools, language, books, teachers, peers, diagrams, calculators, and institutional routines. The relevant contrast is not between pure human cognition and contaminated technological cognition. Salomon et al. (1991) and Perkins (1993) showed that intelligent activity is often distributed across persons and artifacts. The problem is not cognitive distribution itself. The problem emerges when distribution becomes displacement: when tools no longer extend the learner's capacity but substitute for the learner's participation in inquiry, judgment, and responsibility.

Table 1

Distinguishing AI-Enhanced Efficiency From Agency-Preserving Learning

Dimension	AI-enhanced efficiency	Agency-preserving learning
Question formation	The tool generates topics, questions, and directions for the learner.	The learner first articulates problems, uncertainties, and purposes before using AI as a conversational support.
Knowledge construction	The tool produces summaries, explanations, or essays that are accepted as outputs.	AI outputs are treated as provisional materials to be compared, verified, challenged, and revised.
Metacognition	The learner uses AI to reduce effort without monitoring understanding.	The learner uses AI to identify gaps, test comprehension, and plan next steps.
Authorship	Responsibility for claims becomes ambiguous because the output appears ready-made.	The learner records how AI was used and remains accountable for final judgments.
Teacher role	The teacher becomes a detector of misuse or manager of platform access.	The teacher designs tasks, scaffolds judgment, and cultivates responsible inquiry.

The Educational Affordances of Generative AI

A critical account of GenAI should not deny its educational affordances. Several features make it potentially valuable for learning. First, GenAI can provide immediate explanatory support. Students who are hesitant to ask questions in class may use conversational systems to request clarification, examples, analogies, or alternative explanations. This can lower the threshold for participation and support learners who need repeated explanation. Second, GenAI can support formative feedback. It can respond to drafts, identify structural weaknesses, suggest counterarguments, and help students compare different ways of explaining an idea. Feedback is one of the most powerful influences on learning when it clarifies goals, current performance, and next steps (Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006).

Third, GenAI can support differentiation. Students vary in background knowledge, language proficiency, pace, and confidence. Adaptive explanation and low-stakes practice can help some learners enter tasks that would otherwise feel inaccessible. Fourth, GenAI can function as a simulation partner. Students can ask it to play the role of a critic, interviewer, historical figure, patient, client, or peer reviewer. Such uses may expand opportunities for rehearsal, argumentation, and perspective-taking. Fifth, GenAI can help teachers design examples, generate practice materials, and anticipate misconceptions. These teacher-facing affordances are important because the quality of student agency depends partly on the quality of task design and feedback structures.

However, affordances are not outcomes. The existence of a possible educational use does not guarantee educational value. The same capacity that provides immediate explanation can encourage premature closure. The same drafting capacity that helps students overcome a blank page can make it easier to avoid struggling with structure, evidence, and voice. The same personalization that supports access can also narrow learning if students receive frictionless explanations without exposure to disagreement, difficulty, or collective inquiry. Educational value requires pedagogical mediation.

Research on AI in higher education has repeatedly warned that technology-centered approaches often underrepresent pedagogical theory and educator judgment. Zawacki-Richter et al. (2019) found that research on AI applications in higher education had limited critical reflection on risks and weak connections to educational theory. More recent work on AI assistance and student agency has similarly shown that the impact of AI cannot be evaluated only by task performance; it must also be examined in relation to learners' control, self-regulation, and engagement with feedback (Darvishi et al., 2024). This is why the central question is not whether GenAI can help students produce better work, but whether it helps them become better learners.

Three Mechanisms Through Which GenAI May Erode Student Agency

The risks of GenAI are often framed around plagiarism, hallucination, data privacy, or bias. These risks are real and deserve attention (Bender et al., 2021; Floridi & Cowls, 2019; UNESCO, 2023). Yet from an educational perspective, the most serious risks are formative. They concern the kinds of habits and dispositions that repeated tool use may cultivate. Three mechanisms are especially important: the weakening of question formation, the substitution of fluent output for knowledge construction, and the diffusion of responsibility.

The first mechanism is the erosion of question formation. Good learning begins not with answers but with problems. Deweyan inquiry starts from doubt, disturbance, or felt difficulty; knowledge grows through attempts to define and resolve problematic situations (Dewey, 1938). GenAI can invert this sequence. Because students can obtain polished responses quickly, they may come to treat learning as the optimization of prompts rather than the development of questions. Prompting is not the same as questioning. A prompt can be an instruction to the machine; an educational question is an expression of intellectual need, uncertainty, and direction.

When students ask AI to propose research questions before they have encountered a real problem, the learner's own puzzlement may be bypassed.

This does not mean prompt design is trivial. Well-designed prompts can be intellectually demanding. The issue is whether prompt design is grounded in prior inquiry. A student who asks, "What are three weaknesses in my explanation of cognitive load theory?" has already taken a position that can be examined. A student who asks, "Write an essay about cognitive load theory" may receive a complete answer without entering the problem space. The difference is not technical; it is pedagogical. Agency-preserving GenAI use requires students to formulate initial questions, assumptions, and provisional claims before seeking AI assistance.

The second mechanism is the substitution of fluent output for knowledge construction. GenAI systems are powerful producers of plausible language. This creates what may be called a fluency illusion: the appearance of understanding generated by smooth, coherent text. Students may mistake readability for validity, coverage for comprehension, and polish for thought. Research on cognitive offloading shows that people frequently use external tools to reduce cognitive demands (Risko & Gilbert, 2016). Offloading is not inherently harmful; writing notes, drawing diagrams, and using calculators can all support learning. But offloading becomes educationally risky when the learner no longer engages in the core cognitive operations the task was designed to develop.

The problem is intensified in writing. Academic writing is not merely a vehicle for reporting completed thought; it is a method of thinking. Students clarify ideas by organizing paragraphs, searching for evidence, revising claims, and confronting contradictions. If GenAI produces structure, wording, and transitions before students have wrestled with these tasks, writing may become a performance artifact rather than a process of understanding. The educational loss is not only that the student did less work. It is that the student lost the opportunity to form judgment through work.

The third mechanism is the diffusion of responsibility. GenAI complicates authorship. Students may integrate machine-generated phrases, arguments, examples, and citations into their work. Even when no explicit plagiarism occurs, responsibility can become unclear. Who is accountable for a false claim? Who decided that an argument was valid? Who selected the evidence? Responsible agency requires that students stand behind the claims they submit. This does not prohibit assistance. Scholars routinely use dictionaries, editors, peers, databases, and software. The difference is that GenAI can contribute not only language but also reasoning-like structures. Therefore, responsible use requires explicit reflection on how the tool shaped the work and where the learner exercised judgment.

These three mechanisms are mutually reinforcing. If students do not form their own questions, they are more likely to accept AI-generated structures. If they accept generated structures, they are less likely to examine evidence. If they do not examine evidence, responsibility for the final claim becomes thin. Over time, learning may shift from inquiry to curation, from construction to selection, and from responsibility to compliance. This is the central agency risk.

The Paradox of Personalization

Personalization is one of the most attractive promises of AI in education. GenAI can adapt explanations to a student's level, produce examples in familiar domains, translate complex material, and offer unlimited practice. These functions can support access and reduce frustration. However, personalization also contains a paradox. If personalization means that learning environments respond to students in ways that support their growth, it can enhance agency. If personalization means that systems continuously remove difficulty, ambiguity, and friction, it can weaken agency.

Learning requires an appropriate relation between support and struggle. Cognitive load theory reminds educators that unnecessary load can overwhelm learners, especially novices (Sweller, 1988). At the same time, research on guided instruction cautions against assuming that learners benefit from being left entirely to discover complex knowledge without support (Kirschner et al., 2006). The issue is not whether support is good or bad. The issue is whether support is calibrated and gradually withdrawn. A scaffold is educational when it enables learners to do what they could not yet do alone and then fades as competence develops. A substitute is anti-educational when it permanently performs the learner's task.

GenAI personalization easily slides from scaffold to substitute because the tool can keep assisting without natural limits. A human teacher can observe hesitation, press for explanation, ask why, and decide when to withdraw support. AI assistance, unless deliberately designed and pedagogically constrained, may continue to provide answers whenever requested. Students may therefore develop dependence not because they lack ability, but because the learning environment rewards immediate completion more than disciplined independence.

This paradox suggests that agency-preserving education must distinguish between access, assistance, and autonomy. Access means students are not excluded from learning because of language, prior knowledge, disability, or resource gaps. Assistance means they receive support that makes participation possible. Autonomy means they increasingly internalize strategies and judgments. GenAI can contribute to all three, but only if tasks require students to move beyond received output toward explanation, verification, transfer, and self-assessment.

Toward an Agency-Preserving Pedagogy of Generative AI

If GenAI can both support and erode agency, the educational response should not be prohibition alone or adoption alone. The more defensible response is pedagogical design. An agency-preserving pedagogy treats GenAI as a scaffold, interlocutor, and object of critique rather than as a substitute author or automatic authority. Six principles are proposed.

First, problem-posing should precede prompting. Before students use GenAI, they should articulate their own question, confusion, hypothesis, or position. In writing tasks, this may mean submitting a research problem and initial claim before AI consultation. In problem-solving tasks, it may mean predicting a solution strategy before requesting hints. In reading tasks, it may mean identifying unclear concepts before asking for explanation. This principle preserves the learner's ownership of the inquiry. GenAI then responds to a student-initiated problem rather than replacing the problem with a generated task.

Second, AI use should be process-transparent. Students should document when, why, and how they used GenAI. This need not become a bureaucratic burden. A short AI-use log can ask: What did I ask? What did the tool provide? What did I accept, reject, or revise? What evidence did I use to verify the output? Such documentation supports metacognition and responsible agency. It also shifts academic integrity from a purely punitive concern to a learning practice.

Third, students should learn critical verification. GenAI outputs should be treated as claims, not answers. Students should compare outputs with course readings, empirical evidence, peer discussion, and teacher feedback. This is especially important because large language models can generate inaccurate or fabricated information while maintaining confident style (Bender et al., 2021; Kasneci et al., 2023). Verification tasks can include source checking, argument mapping, error analysis, and comparison of multiple AI-generated responses. The goal is not to teach students distrust for its own sake, but to cultivate judgment.

Fourth, AI assistance should be scaffolded and faded. In early stages, students may use GenAI for vocabulary clarification, example generation, or low-stakes practice. As competence develops, tasks should require more independent planning, drafting, and justification. Teachers can create staged assignments in which AI is permitted for some phases and restricted in others. For example, AI may be used to critique a student's outline but not to write the first draft; or it may be used after an initial in-class attempt, not before. The sequencing matters. Agency grows when learners experience assistance as a means to independence.

Fifth, assessment should move from product-only evaluation to process-and-judgment evaluation. If grades focus only on final polished outputs, GenAI will predictably be used to optimize products. Assessment should therefore include oral defense, draft histories, reflective memos, annotated bibliographies, in-class synthesis, peer discussion, and explanation of revisions. These practices do not eliminate AI use. They make visible the learner's understanding, decisions, and responsibility. Assessment should ask not only "What did the student submit?" but also "How did the student inquire, judge, revise, and justify?"

Sixth, GenAI should become an object of inquiry, not only a learning aid. Students should examine how AI responses are generated, why outputs may be biased or incomplete, and how design choices influence knowledge practices. This aligns with emerging AI competency frameworks that emphasize responsible use, human judgment, and critical understanding rather than mere operational skill (UNESCO, 2024a, 2024b). When students study the tool itself, they are less likely to treat it as an invisible authority. They learn to see AI as a fallible participant in a broader ecology of knowledge.

Table 2

Agency Risks and Pedagogical Responses in GenAI-Supported Learning

Agency risk	Typical classroom manifestation	Pedagogical response
Question displacement	Students ask AI to generate topics or thesis statements before identifying their own problem.	Require a pre-AI problem statement, hypothesis, or question memo.
Fluency illusion	Students accept coherent AI text as evidence of understanding.	Use verification tasks, error analysis, and comparison with course sources.
Cognitive over-offloading	Students use AI to plan, draft, explain, and revise without doing the core intellectual work.	Sequence AI access after initial independent attempts and fade support across tasks.
Responsibility diffusion	Students cannot explain why claims are included or how AI shaped the final work.	Require AI-use logs, reflective memos, and oral defense of key claims.
Assessment distortion	Grades reward polished products more than inquiry and judgment.	Assess process, reasoning, revision history, and transfer across contexts.
Teacher role reduction	Teachers become primarily detectors of AI misuse.	Position teachers as designers of inquiry, feedback, and ethical judgment.

The Teacher's Role: From Detection to Pedagogical Judgment

The rise of GenAI has placed teachers in a difficult position. They are expected to prevent misuse, redesign assessment, teach AI literacy, support learning, and maintain fairness, often without sufficient time or institutional support. A narrow response turns teachers into detectors of misconduct. This role is educationally inadequate. Detection may be necessary in some contexts, but it cannot be the foundation of a GenAI pedagogy.

An agency-preserving approach positions teachers as designers of conditions for judgment. Teachers decide which tasks require unaided effort, which tasks benefit from AI-supported exploration, and which tasks should explicitly analyze AI output. They help students understand when assistance is appropriate, how to verify claims, and why responsibility cannot be delegated to a tool. They also maintain the human dimensions of education that AI cannot replace: interpreting students' developing understanding, responding to emotional and social contexts, modeling intellectual humility, and inviting learners into communities of inquiry.

This teacher role is consistent with the broader tradition of scaffolding. In Vygotskian terms, learning occurs through mediated participation in activities that learners gradually appropriate (Vygotsky, 1978). In cognitive apprenticeship, expert processes are modeled, coached, and gradually transferred to learners (Collins et al., 1989). GenAI can participate in this ecology, but it cannot replace pedagogical judgment. The teacher must decide how AI assistance is related to the educational purpose of the task. Without this judgment, AI use is likely to be governed by convenience rather than formation.

Teachers also need institutional protection from contradictory expectations. Institutions cannot simultaneously demand strict prohibition, encourage innovation, maintain traditional assessment unchanged, and provide no time for redesign. Faculty development should therefore move beyond tool demonstrations. Teachers need support in designing AI-integrated assignments, evaluating process evidence, discussing academic integrity constructively, and addressing uneven student access. The goal is not to make every teacher a technical expert. It is to support teachers as educational professionals capable of making defensible judgments about when, how, and why GenAI should be used.

Assessment Redesign and the Recovery of Responsibility

Assessment is where the tension between efficiency and agency becomes most visible. When assessment rewards final products that can be generated or heavily improved by AI, students have strong incentives to outsource. This does not mean that all traditional writing assignments are obsolete. It means that product-only assessment is no longer sufficient as evidence of learning. The educational response should be to recover responsibility by making thinking processes more visible.

Several assessment practices are useful. First, staged assignments can require students to submit research questions, annotated sources, outlines, draft excerpts, peer feedback, AI-use logs, and final reflections. Second, oral or written defenses can ask students to explain why they made specific choices. Third, in-class writing or problem-solving can provide evidence of unaided competence. Fourth, revision-based assessment can reward the quality of improvement rather than only the quality of final prose. Fifth, comparative analysis tasks can ask students to critique AI-generated responses using course concepts. These practices do not merely police AI. They create conditions in which students must demonstrate agency.

Assessment redesign also requires clearer distinctions among assistance, collaboration, and authorship. Students should know whether AI may be used for brainstorming, grammar correction, feedback, coding support, translation, source discovery, or drafting. These categories should not be treated as morally identical. For example, using AI to generate possible counterarguments for critique differs from submitting AI-generated analysis as one's own reasoning. A mature academic integrity framework should specify permitted, limited, and prohibited uses in relation to the learning outcomes of each task.

The principle is simple: the closer AI assistance comes to performing the central learning outcome, the stronger the need for transparency, restriction, or redesign. If the outcome is grammar accuracy, AI proofreading may perform the outcome. If the outcome is conceptual argumentation, AI-generated argument structures may perform the outcome. If the outcome is critical evaluation of AI, then AI output becomes appropriate evidence for analysis. The same tool use can therefore be acceptable in one task and unacceptable in another. This is not inconsistency; it is alignment between means and educational purpose.

Curriculum Implications: AI Literacy as Educational Formation

AI literacy is often interpreted as knowing how to use tools effectively. That interpretation is too narrow. In an agency-preserving curriculum, AI literacy includes technical, epistemic, ethical, and metacognitive dimensions. Students should understand what GenAI can do, but also how it can mislead; how to prompt, but also when not to prompt; how to use assistance, but also how to remain responsible for judgment.

A curriculum for student agency should include at least four components. The first is functional understanding: students learn the basic capacities and limitations of GenAI systems, including their dependence on training data and probabilistic text generation. The second is epistemic evaluation: students learn to verify claims, identify unsupported assertions, compare sources, and recognize the difference between plausible language and justified knowledge. The third is metacognitive regulation: students learn to decide when AI use supports their learning and when it reduces necessary struggle. The fourth is ethical responsibility: students learn norms of disclosure, authorship, fairness, and respect for others' intellectual labor.

These components should not be isolated in a single technology module. They should be integrated across disciplines. In writing courses, students can critique AI-generated essays for evidence and voice. In science courses, they can test AI explanations against empirical models. In teacher education, they can design lessons that use AI as a scaffold while protecting student reasoning. In humanities courses, they can examine how language, authorship, and interpretation change when texts are machine-generated. The aim is to form learners who can use AI without surrendering inquiry.

This approach also avoids a simplistic digital native assumption. Students may use AI frequently without understanding it critically. Familiarity is not literacy. A student can be skilled at generating outputs and weak at evaluating them. Therefore, AI literacy must be linked to the broader educational goals of critical thinking, self-regulated learning, disciplinary reasoning, and responsible participation in knowledge communities.

Equity, Access, and the Uneven Distribution of Agency

Although this article avoids a policy-centered argument, equity remains educationally relevant. GenAI may reduce some learning barriers while intensifying others. Students with limited access to tutoring, language support, or academic feedback may benefit substantially from AI assistance. At the same time, students with stronger prior knowledge may be better positioned to evaluate AI outputs, detect errors, and use assistance strategically. Without pedagogical guidance, GenAI could therefore widen differences in agency rather than simply democratize learning.

The key issue is not access to the tool alone. Access without critical capacity can produce dependence. Access with guided inquiry can produce empowerment. Students who already possess strong disciplinary understanding can use GenAI to extend thinking; students with weaker foundations may accept fluent errors. This suggests that educators should not assume that open access automatically equals educational fairness. Fairness requires instruction in verification, transparent norms, and task designs that support all learners in becoming more capable users.

There is also a risk that institutions with fewer resources will adopt AI as a substitute for human teaching, while more privileged contexts use it as an enrichment tool under expert guidance. Such differentiation would be educationally harmful. The agency-preserving position is clear: GenAI should augment human teaching and learner inquiry, not replace the relational and professional work through which students become responsible knowers. Efficiency should not become an excuse for thinning the educational encounter.

A Normative Framework for Responsible Educational Use

The preceding analysis can be summarized in a normative framework. GenAI use is educationally defensible when it satisfies three conditions. First, it preserves the learner's ownership of the problem. The student should be able to explain what question is being pursued and why it matters. Second, it preserves the learner's participation in knowledge construction. The student should engage in interpretation, evidence evaluation, and revision rather than merely selecting from generated outputs. Third, it preserves the learner's responsibility for judgment. The student should be able to disclose assistance, justify claims, and accept accountability for the final work.

These conditions can guide teachers and institutions more effectively than blanket rules. A ban may be appropriate for tasks designed to assess unaided performance. Open use may be appropriate for tasks designed to teach critique of AI outputs. Limited use may be appropriate for drafting, feedback, or language support. The educational question is always: Does this use strengthen or weaken the capabilities the task is meant to cultivate?

This framework also rejects the assumption that AI use is automatically innovative. Innovation in education is not the introduction of a new tool; it is the improvement of learning conditions in relation to worthwhile aims. A classroom saturated with GenAI but poor in questioning, dialogue, and judgment is not advanced. A classroom that uses GenAI sparingly but thoughtfully to sharpen inquiry may be more genuinely innovative. Educational progress should be measured by the growth of student agency, not by the quantity of technological integration.

Discussion

This article has argued that GenAI should be understood through the lens of student agency. The central issue is not whether AI can generate useful outputs, because it can. Nor is the central issue whether misuse is possible, because it is. The deeper question is whether educational practice can integrate GenAI in ways that preserve students as questioners, knowledge builders, self-regulated learners, and responsible authors.

The analysis contributes to educational theory in three ways. First, it reframes GenAI from a tool-adoption problem to a subject-formation problem. This reframing prevents educators from equating efficiency with learning. Second, it distinguishes between cognitive distribution and cognitive displacement. Learning has always depended on tools and social mediation, but tools become educationally problematic when they replace the learner's participation in the central operations of inquiry. Third, it offers a practical framework for pedagogy: problem-posing before prompting, process transparency, critical verification, scaffolded assistance, assessment redesign, and teacher-guided ethical reflection.

The implications for research are substantial. Empirical studies should examine not only achievement outcomes or attitudes toward AI, but also changes in question quality, self-regulation, epistemic judgment, authorship awareness, and transfer of learning. Researchers should ask whether AI-supported students can perform independently after scaffolds are withdrawn. They should examine whether different forms of AI feedback enhance or reduce learner control. They should also investigate disciplinary differences. The agency risks in writing, programming, mathematics, language learning, and clinical reasoning may not be identical.

The implications for practice are equally clear. Teachers should not respond to GenAI by simply adding detection software or by allowing unrestricted use. They should redesign tasks so that students must show their reasoning, justify their choices, and reflect on AI assistance. Institutions should support this redesign rather than placing the burden entirely on individual teachers. Students should be taught that responsible AI use is not a matter of hiding or displaying tool use, but of remaining intellectually accountable for learning.

This article has limitations. It is a theoretical analysis rather than an empirical study. Its claims should be tested in classroom research across educational levels, disciplines, and cultural contexts. It also treats GenAI broadly, though specific systems differ in design, reliability, transparency, and educational affordances. Finally, the concept of agency itself requires careful operationalization if it is to guide empirical assessment. Future research should develop valid measures of student agency in AI-mediated learning and examine how pedagogical designs affect agency over time.

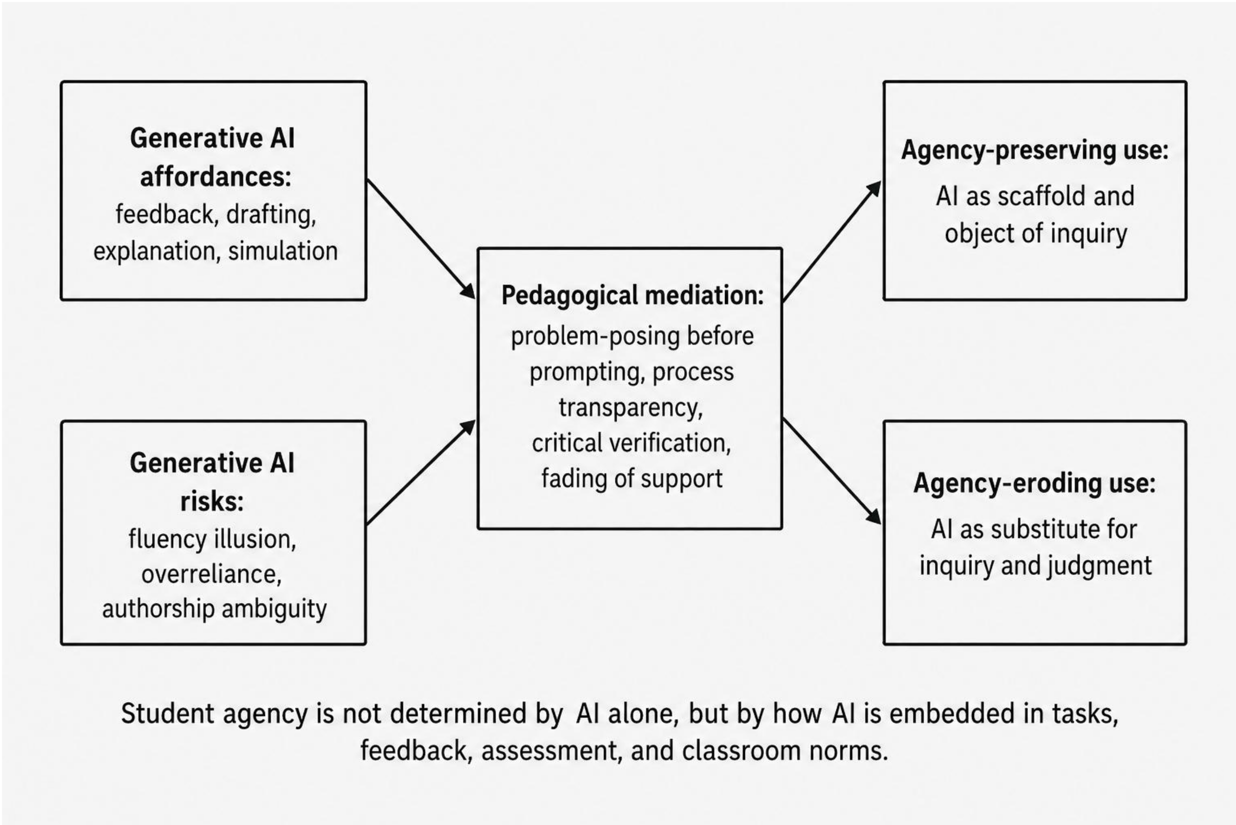
Conclusion

Generative AI forces education to clarify what it means by learning. If learning is defined as the production of correct or polished outputs, then GenAI will appear mainly as a powerful productivity tool. If learning is understood as the formation of questioning, understanding, judgment, and responsibility, then GenAI becomes more ambiguous. It can support learning when it functions as a scaffold for inquiry; it can erode learning when it substitutes for the learner's own intellectual work.

The educational task is therefore not to defend an obsolete model of unaided learning, nor to surrender education to technological efficiency. The task is to design learning environments in which students use GenAI without losing themselves as agents. Students should learn to ask before prompting, think before accepting, verify before citing, revise before submitting, and take responsibility before claiming authorship. In this sense, the future of education with GenAI depends less on the intelligence of machines than on the seriousness with which educators protect and cultivate the agency of learners.

Figure 1

A Pedagogical Framework for Agency-Preserving GenAI Use



Note. The framework emphasizes that GenAI does not determine learning outcomes by itself. Student agency depends on pedagogical mediation, task design, assessment, and classroom norms.

Declarations

Funding: The author received no financial support for the authorship or publication of this article.

Conflict of Interest: The author declares no conflict of interest.

Data Availability: No datasets were generated or analyzed for this article.

Author Contributions: [Author Name] conceptualized the article, conducted the literature-based theoretical analysis, and wrote the manuscript.

References

Baker, T., & Smith, L. (2019). Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges. Nesta.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. Annual Review of Psychology, 52(1), 1-26. https://doi.org/10.1146/annurev.psych.52.1.1

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610-623. https://doi.org/10.1145/3442188.3445922

Bereiter, C. (2002). Education and mind in the knowledge age. Lawrence Erlbaum Associates.

Biesta, G. (2015). What is education for? On good education, teacher judgement, and educational professionalism. European Journal of Education, 50(1), 75-87. https://doi.org/10.1111/ejed.12109

Biesta, G. (2017). The rediscovery of teaching. Routledge.

Bruner, J. S. (1960). The process of education. Harvard University Press.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), Knowing, learning, and instruction: Essays in honor of Robert Glaser (pp. 453-494). Lawrence Erlbaum Associates.

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

Dewey, J. (1916). Democracy and education: An introduction to the philosophy of education. Macmillan.

Dewey, J. (1938). Experience and education. Macmillan.

Emirbayer, M., & Mische, A. (1998). What is agency? American Journal of Sociology, 103(4), 962-1023. https://doi.org/10.1086/231294

Fawns, T. (2022). An entangled pedagogy: Looking beyond the pedagogy-technology dichotomy. Postdigital Science and Education, 4, 711-728. https://doi.org/10.1007/s42438-022-00302-7

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. Harvard Data Science Review, 1(1). https://doi.org/10.1162/99608f92.8cd550d1

Freire, P. (1970). Pedagogy of the oppressed. Continuum.

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

Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign.

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. Educational Psychologist, 41(2), 75-86. https://doi.org/10.1207/s15326985ep4102_1

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed: An argument for AI in education. Pearson.

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

Perkins, D. N. (1993). Person-plus: A distributed view of thinking and learning. In G. Salomon (Ed.), Distributed cognitions: Psychological and educational considerations (pp. 88-110). Cambridge University Press.

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. Trends in Cognitive Sciences, 20(9), 676-688. https://doi.org/10.1016/j.tics.2016.07.002

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55(1), 68-78. https://doi.org/10.1037/0003-066X.55.1.68

Salomon, G., Perkins, D. N., & Globerson, T. (1991). Partners in cognition: Extending human intelligence with intelligent technologies. Educational Researcher, 20(3), 2-9. https://doi.org/10.3102/0013189X020003002

Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), The Cambridge handbook of the learning sciences (pp. 97-118). Cambridge University Press.

Selwyn, N. (2019). Should robots replace teachers? AI and the future of education. Polity Press.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257-285. https://doi.org/10.1207/s15516709cog1202_4

UNESCO. (2023). Guidance for generative AI in education and research. UNESCO.

UNESCO. (2024a). AI competency framework for students. UNESCO.

UNESCO. (2024b). AI competency framework for teachers. UNESCO.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. Learning, Media and Technology, 45(3), 223-235. https://doi.org/10.1080/17439884.2020.1798995

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), Metacognition in educational theory and practice (pp. 277-304). Lawrence Erlbaum Associates.

Zawacki-Richter, O., Marin, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? International Journal of Educational Technology in Higher Education, 16, Article 39. https://doi.org/10.1186/s41239-019-0171-0

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. Theory Into Practice, 41(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2