



Embodied Competence in the AI Era: Must Educators Personally Practice the Human Operating System to Cultivate It?

Executive Summary

Framing the central dilemma

In AI-saturated classrooms, the most consequential educational question is no longer "How do students get correct answers?" but "What human capacities remain *educationally non-negotiable* when answers are cheap, fluent, and sometimes wrong?" Generative AI systems can produce plausible explanations, essays, and arguments at low friction, but they also introduce predictable epistemic hazards: over-trust in confident outputs, "automation bias," and skill atrophy when learners (or educators) rely on the system as a cognitive crutch rather than as a tool for thought. ¹

The problem posed in this report is therefore not primarily about technical AI literacy. It is about **epistemic embodiment**: whether teachers must *personally enact* (and be seen to enact) five "human operating system" competencies so that students can develop them with authenticity, transfer, and moral legitimacy under conditions of pervasive AI assistance.

Bottom line answer: A is necessary but rarely sufficient; B is often minimally required

Across learning sciences, cognitive psychology, and epistemology of education, **conceptual understanding + activity design (A) is necessary but tends to be insufficient** for the five competencies when (i) the competencies are *performative and dispositional* (requiring judgment under uncertainty), (ii) transfer is a goal, and (iii) classroom norms must be stable under the "epistemic disruptor" introduced by AI. ²

Instead, the best-supported inference is a **qualified version of B**:

- Teachers do **not** need to be "AI power users," but they do need **credible, practiced competence** in (a) *making epistemic moves visible* (how to question, verify, and revise), (b) managing the *temptation to offload cognition*, and (c) modeling ethical responsibility and agency when AI could easily absorb authorship and judgment. ³
- Where teacher embodiment matters most is **not** the accuracy of the teacher's content recall; it is the teacher's *epistemic legitimacy* as a guide for what counts as warranted belief, responsible use, and human authorship in an environment where a machine can sound authoritative. ⁴

Why embodiment becomes more load-bearing under generative AI

Three converging literatures make teacher embodiment unusually consequential in AI-mediated learning:

1. **Cognitive apprenticeship and observational learning** show that novices learn complex cognition partly by seeing experts *perform and articulate* the invisible parts: goal setting, noticing errors, deciding what to trust, and self-correcting. The cognitive apprenticeship tradition explicitly emphasizes “making thinking visible” through modeling, coaching, and scaffolding. 5
2. **Transfer and metacognition research** suggests that far transfer requires learners to acquire not only knowledge and procedures but also *metacognitive monitoring* and *dispositions* to apply skills without prompting—precisely the features that are easiest to short-circuit via AI offloading. 6
3. New **2023-2026 empirical evidence on AI use** indicates that AI assistance can improve immediate performance while undermining agency or subsequent independent performance when the tool is withdrawn, unless guardrails and reflective practices are designed into the learning environment. 7

Competency-by-competency headline conclusions

Below, each of the five competencies is evaluated against A vs B, with attention to mechanism, evidence strength, and AI-specific affordances/risks.

Question Formulation / Inquiry Discipline

A can produce *structured participation* (students do inquiry tasks), but B is typically required for *inquiry as a habit of mind* because question quality emerges from iterative refinement, noticing confusion, and epistemic goal-setting. Prompting with LLMs converts question discipline into an interactive, iterative practice, increasing the value of teachers who can model iteration and constraint-setting (not merely assign it). 8

Critical Evaluation of AI Outputs

B is the strongest requirement here. Verification, triangulation, and bias detection are not “content” to be delivered; they are *epistemic actions* under uncertainty. Recent intervention evidence shows verification behavior can be shifted (e.g., inoculation-style prompts), but effective classroom cultivation depends on teachers who themselves practice detecting failure modes and demonstrating lateral reading / cross-checking, especially under time pressure. 9

Ethical Judgment and Responsible Use

A can transmit rules, but B is often required for legitimacy: students calibrate norms via the hidden curriculum (what adults *actually do*). Meta-analytic evidence that teacher credibility relates meaningfully to student outcomes underscores that trustworthiness and caring—dimensions tightly linked to perceived authenticity—shape receptivity. AI introduces new moral objects (authorship, attribution, privacy, bias), which are hard to teach credibly without teachers engaging in the dilemmas themselves. 10

Adaptive Learning Under Technological Uncertainty

B is moderately required: adaptive learning is largely meta-learning (monitoring what one knows, updating strategies, and maintaining epistemic humility). Teacher embodiment matters because

teachers are the system's primary "meta-model": students observe whether uncertainty leads to inquiry and revision or to avoidance and brittle certainty. Recent teacher readiness evidence (2026) indicates adoption involves perceived opportunities and threats, classroom-use confidence, and ethical caution—suggesting that "methodological add-ons" without personal adaptation are unstable. ¹¹

Keeping the Human Above the Machine

B is strongly required because the competency is fundamentally normative: it concerns agency, responsibility, and human judgment. Empirical work on AI assistance and agency shows learners can become reliant rather than learning from AI prompts, and self-regulated replacements are less effective than the AI itself—implying a real risk that the machine becomes the de facto regulator unless teachers actively model human-centered control. ¹²

Executive synthesis: "minimum viable embodiment" rather than maximal purity

A productive way to resolve the A vs B dichotomy is to treat embodiment as a **continuum of instructional legitimacy**:

- **Design-only competence** (A) can sometimes work for bounded procedural skills and short-term outcomes.
- **Minimum viable embodiment** is required when instruction aims for (i) transfer, (ii) stable epistemic norms, and (iii) learner agency under AI temptation.
- "Embodiment" should be defined pragmatically as **teachers' repeated, demonstrable practice of the target epistemic and ethical actions in authentic AI-mediated contexts**, including visible self-correction and explicit justification. ¹³

The remainder of this report builds a rigorous argument for that position, distinguishing **well-supported claims** (meta-analytic and widely replicated), **emerging claims** (recent studies; some preprints), and **speculative claims** (plausible but under-validated, especially regarding long-term cognitive "atrophy").

Conceptual clarification of embodied competence

What "embodied competence" means in this report

This report defines **embodied competence** as *performed, situated, and publicly visible enactment* of a competency in the same class of environments where students must learn to enact it (here: environments in which generative AI is available and tempting). This is close to an "epistemology of practice" orientation associated with Donald Schön ¹⁴, where professionals develop knowledge-in-action and reflection-in-action under conditions of uncertainty rather than merely applying explicit rules. ¹⁵

In educational terms, "embodiment" is not mysticism or personality. It is the **instructionally relevant portion** of a teacher's practice that students can perceive and internalize:

- visible reasoning (why this claim is warranted),
- visible checking (how to test it),
- visible humility (what I don't know and how I proceed anyway),
- visible ethics (how I decide what is permissible and why),
- visible agency (where the human remains responsible). ¹⁶

Philosophical grounding: experience, dialogue, and reflective practice

Three canonical philosophical lenses converge on the centrality of lived practice:

- John Dewey ¹⁷ argues that education is rooted in experience and that learning quality depends on how experiences are structured and reflected upon, rather than on verbal transmission alone. ¹⁸
- Paulo Freire ¹⁹ emphasizes dialogical, problem-posing pedagogy and critiques “banking” education; in AI contexts, this is directly relevant because generative AI can re-institute a new “banking” dynamic (answers deposited into students) unless pedagogy keeps knowledge as a co-constructed, critically examined object. ²⁰
- Schön’s reflective practitioner frame treats uncertainty as the normal condition of professional action, implying that teaching epistemic judgment under AI requires educators to enact reflective inquiry rather than pretend certainty. ²¹

Learning-science grounding: cognitive apprenticeship and observational learning

Cognitive apprenticeship theory—associated with Allan Collins ²² and colleagues—argues that expert cognitive processes are often invisible and must be externalized via modeling, coaching, scaffolding, articulation, reflection, and exploration. ²³

A complementary mechanism is observational learning: classic evidence (reviewed in modern scholarship) shows that learners acquire behaviors by observing models and the consequences of modeled behavior. ²⁴ This observational pathway becomes more important when the competency is partly normative (what one *ought* to do) and partly procedural (how one *actually* does it), as with verification, ethical restraint, and intellectual ownership.

Review of evidence on modeling vs facilitation

Well-supported evidence: guidance, modeling, and making thinking visible outperform “design-only” approaches for complex skills

Meta-analytic and synthesis evidence supports three claims that matter directly for A vs B:

First, **guided discovery and scaffolding** are generally more effective than minimally guided approaches, especially for novices. A major meta-analysis on discovery-based instruction finds that *guided* discovery improves learning outcomes substantially more than unguided discovery, implying that “facilitation” must include skillful guidance rather than mere activity assignment. ²⁵

Second, **inquiry learning benefits from instructional guidance**. Meta-analytic reviews of inquiry-based instruction show positive effects when inquiry is structured and supported rather than left to pure exploration. ²⁶ This implies a teacher role that is not only architectural (designing inquiry tasks) but also *performative* (modeling how inquiry proceeds when it meets ambiguity).

Third, **critical thinking can be taught, but transfer depends on dispositions and metacognitive monitoring**. A major meta-analysis finds that instruction can improve critical thinking outcomes, but effect sizes vary widely depending on how instruction is designed and enacted. ²⁷ Transfer-focused theory argues that instruction must cultivate (a) dispositions, (b) skills, (c) structural understanding, and (d) metacognitive monitoring—the last two being hard to teach credibly without demonstration of self-checking in real time. ²⁸

Modeling as a mechanism: examples, worked solutions, and explanations

Beyond the apprenticeship tradition, example-based learning provides empirically supported insight into why modeling matters. A modern meta-analysis of the worked example effect (focused on rigorous designs) supports the conclusion that learners benefit when they study worked solutions rather than only solving problems unaided—especially under conditions of cognitive load.²⁹ A meta-analytic review also finds that **instructional explanations accompanying examples** improve learning, reinforcing that the “why” and “how” of expert reasoning—not just exposure to tasks—drives learning gains.³⁰

In AI contexts, this matters because the “worked example” is now often produced by the model; without teacher modeling of *how to interrogate* the example, students may learn to accept fluent outputs rather than develop evaluative competence.

Transfer logic: why teacher embodiment becomes more important when the goal is “use without prompting”

Transfer scholarship emphasizes that transfer is fostered when instruction both “hugs” (resembles target contexts) and “bridges” (explicitly connects principles across contexts).³¹

Generative AI changes the target context: learners will face settings where AI is available, persuasive, and fallible. Therefore, the “hugging” component of transfer now plausibly requires that teachers **model the target performance under AI presence** (e.g., verifying an AI answer under time pressure). Without such modeling, students may complete classroom activities but fail to generalize the relevant epistemic actions when AI is present outside the classroom.³²

AI as epistemic disruptor and implications for teacher legitimacy

Generative AI destabilizes “teacher as knowledge source,” intensifies “teacher as epistemic mentor”

Recent work frames generative AI as introducing a “crisis of epistemic authority,” because systems can produce fluent, confident outputs that learners may treat as authoritative—often without adequate reflection on limitations, bias, or verification.³³ In such contexts, the teacher’s comparative advantage shifts away from being a repository of information and toward being an **epistemic mentor**: someone who teaches how to justify, check, and responsibly use knowledge claims.

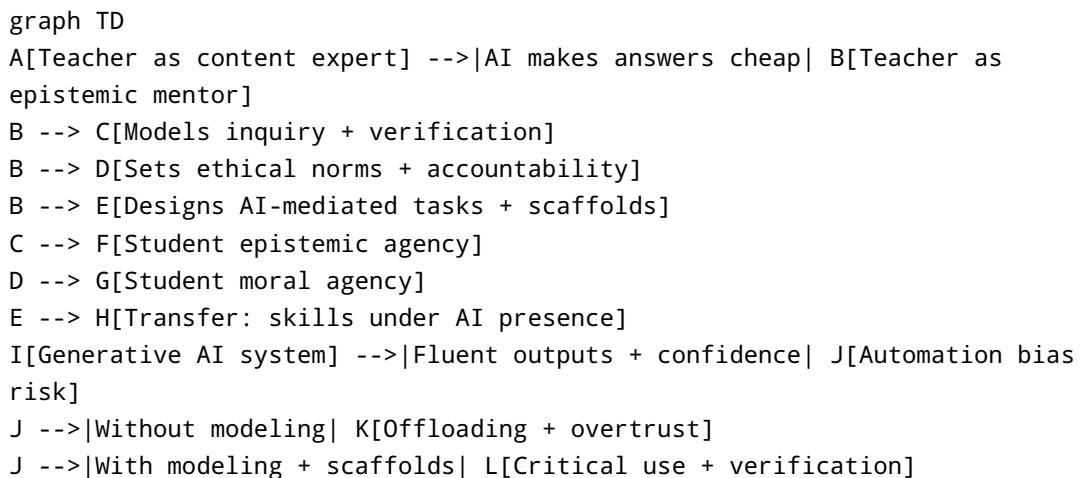
This is consistent with a broader argument about teacher knowledge in the age of generative AI, articulated in work by Punya Mishra³⁴ and colleagues, emphasizing that teacher competencies must evolve to respond to generative AI’s distinctive attributes.³⁵

Epistemic legitimacy, credibility, and the hidden curriculum

If generative AI shifts authority, teacher legitimacy becomes less about being correct and more about being **credible** as a guide for good epistemic conduct. A meta-analysis of teacher credibility finds a meaningful relationship between perceived teacher credibility (competence, trustworthiness, caring) and student outcomes.³⁶ This supports the claim that student receptivity is shaped by perceived authenticity and trust, not only by lesson design.

Simultaneously, the hidden curriculum literature emphasizes that students learn implicit messages about norms, values, and behavior through daily educational processes—not only through explicit instruction.³⁷ In AI contexts, a teacher who assigns verification tasks but never demonstrates verification may inadvertently teach: “verification is school theater, not a real norm.” Conversely, a teacher who openly cross-checks AI outputs and narrates uncertainty communicates that epistemic responsibility is central, not ornamental.³⁸

Visualizing the role shift



Empirical findings from 2023 to 2026 on generative AI, critical thinking, and cognitive offloading

Evidence cluster: AI assistance can raise immediate performance while weakening agency or independent capability

A central empirical pattern in the emerging literature is **performance/learning divergence**: assistance can improve short-term outcomes while harming longer-term independent competence if learners offload core cognitive work.

A large randomized controlled experiment in a peer-feedback context ($N \approx 1,625$ across 10 courses) found that students tended to **rely on rather than learn from** AI assistance; when AI prompts were removed, self-regulated strategies helped but were not as effective as AI support.³⁹ This matters for the “human OS” because it implies that without explicit cultivation of self-regulation and evaluative judgment, AI may become the hidden executive function of the learning process.

Correlational evidence aligns with this: a 2025 study of university students found that higher AI dependence was associated with **lower critical thinking**, with cognitive fatigue partially mediating the relationship; information literacy buffered some negative effects but could also amplify fatigue under high AI reliance.⁴⁰

A large survey study of 319 knowledge workers reports that confidence in AI for the task predicts reduced enactment of critical thinking, while task-specific self-confidence predicts greater critical thinking enactment—supporting the mechanism that over-trust can narrow human cognitive engagement.⁴¹

Finally, a prominent 2024 field experiment preprint in high-school math contexts reports that access to a GPT-4 tutor improved performance, but when access was removed, students performed worse than those who never had access—while guardrails designed to protect learning reduced negative effects.

⁴² This result is not yet as settled as peer-reviewed work, but it coheres with the agency RCT and dependence findings, and it directly implicates the need for teachers to manage offloading through pedagogy rather than policy alone. ⁴³

Evidence cluster: verification behaviors can be trained, but instruction must confront time pressure and realistic AI use

A 2026 experimental study on inoculation training finds that generic forewarning messages increased verification behaviors for ChatGPT-generated academic summaries, while time pressure reduced fact-checking—highlighting the speed-accuracy trade-off that classroom instruction must address explicitly.

⁴⁴

Parallel evidence from the digital information evaluation literature shows that learners commonly struggle to evaluate credibility without targeted instruction, and interventions that teach lateral reading have shown improvements across diverse populations though measurement and scaling challenges remain. ⁴⁵

Because generative AI can produce fluent misinformation (“hallucinations”), the epistemic burden shifts toward verification skills. Surveys of hallucination phenomena in LLMs emphasize that models can generate plausible but nonfactual content and that detection/mitigation remains an active research area—meaning educators should treat error as a normal condition rather than as an edge case. ⁴⁶

Evidence cluster: teachers' readiness and competence are uneven, implying embodiment cannot be assumed

A 2026 jurisdiction-wide survey of teachers (Australian Capital Territory) identifies readiness dimensions that mix perceived opportunities (e.g., workload reduction) with perceived threats (e.g., risks to student autonomy, social-emotional learning, and teacher roles), and suggests limited clarity of subjective norms due to weak institutional guidance. ⁴⁷

Systematic review evidence also finds that teacher-focused PD is underrepresented relative to AI-in-teaching research, reinforcing the concern that teachers are asked to manage an epistemic disruptor without adequate structured preparation. ⁴⁸

Timeline of the densest evidence cluster

```
timeline
  title Evidence relevant to embodiment in GenAI classrooms (2023-2026)
  2023 : Teacher knowledge frameworks for GenAI (TPACK-inspired) emphasize new teacher competencies
  2024 : RCT shows AI assistance can reduce student agency when learners rely on support
  2024 : Review work consolidates evidence on teaching lateral reading (verification routines)
  2024 : Field experiment preprint suggests GPT tutoring can harm learning when withdrawn unless guardrails exist
```

2025 : Survey data: GenAI confidence associated with reduced critical thinking enactment in knowledge work

2025 : Empirical study links AI dependence with lower critical thinking via fatigue (moderated by information literacy)

2026 : Inoculation training increases verification behaviors; time pressure reduces checking

2026 : Teacher readiness survey identifies mixed enthusiasm/caution and limited normative guidance

Case studies where teachers practiced before teaching

The studies below are selected because they contain *observable educator practice* (teachers or preservice teachers actively using GenAI, reflecting, revising, and encountering failure modes), not only policy guidance. They therefore illuminate what “practice-before-teaching” can look like operationally.

Case study synthesis table

Case	Context and design	What “practice” consisted of	Main findings relevant to embodiment	Evidence strength (for this report’s question)
Preservice teachers using ChatGPT for lesson planning (2024)	Qualitative study in a methods course; preservice secondary math teachers used ChatGPT for solving a math problem and supporting lesson planning & reflection. ⁴⁹	Iterative use of ChatGPT; compare own plan vs AI plan; reflective critique of AI outputs. ⁴⁹	Participants often recognized pedagogical weaknesses (generic, teacher-centered), but were less accurate evaluating mathematical correctness, sometimes re-labeling wrong answers as “different approaches.” This is a direct argument for teacher-modeled verification routines before asking students to evaluate AI. ⁴⁹	Strong for “evaluation requires practice”: shows common miscalibration and the need for modeled checking.

Case	Context and design	What “practice” consisted of	Main findings relevant to embodiment	Evidence strength (for this report’s question)
ChatGPT-supported lesson plan critiques (2025)	Quasi-experiment (N=48 preservice teachers) comparing ChatGPT-supported critique vs control; used cognitive/metacognitive/affective guiding questions; analyzed with epistemic network analysis. 50	Practice was structured critique and revision with/without ChatGPT scaffolds. 50	ChatGPT improved cognitive and affective critiques but not metacognitive critiques— suggesting that metacognitive regulation is not “automatically” learned via AI scaffolds and may require explicit modeling of self-monitoring by educators. 50	Strong for “design plus modeling”: AI scaffolds help some dimensions, not metacognitive ones.
Teachers using ChatGPT for self-directed professional development (2025)	Mixed-methods: 298 teacher surveys + interviews + chat log documents; focus on how teachers used ChatGPT for PD. 51	Teachers used ChatGPT to generate lesson ideas/structures; monitored goals/time; reflected on outputs. 51	Teachers primarily used ChatGPT for inspiration and structuring materials; the study reports evidence of self-management and self-monitoring capabilities during AI-supported PD—an existence proof that “teacher embodied practice” can be operationalized as reflective AI use, not just tool adoption. 51	Moderate-strong for feasibility of embodied practice as PD; less direct on student outcomes.

Case	Context and design	What “practice” consisted of	Main findings relevant to embodiment	Evidence strength (for this report’s question)
GenAI in teacher education—preparedness gap (2025)	Exploratory case study: 52 preservice teachers + 21 teacher educators (surveys and focus groups). 52	Practice was collective exploration and articulation of barriers/facilitators. 52	Reports moderate digital literacy but low AI literacy; preservice teachers not prepared for GenAI prevalence; highlights institutional need for PD that includes hands-on experience rather than conceptual briefings. 52	Moderate: good for diagnosing need; not a causal test of embodiment.
Teacher readiness for GenAI adoption (2026)	Cross-sectional survey (N=111 teachers) using Theory of Planned Behaviour lens; PCA yields readiness dimensions. 47	Practice captured as self-reported experiences, skills, concerns with GenAI. 47	Readiness clusters include “integration skills” and strong ethical/pedagogical caution; subjective norms did not emerge, suggesting teachers lack shared practice norms—reinforcing the need for embodied, community-based PD (practice with shared standards). 47	Moderate: informs PD design and why embodiment may not emerge spontaneously.
Verification behavior can be trained (2026)	Experimental inoculation study: inoculation message vs none; examines verification intentions and behaviors; time pressure included. 44	Practice was prompted verification during AI summary task. 44	Inoculation improved verification behavior but not intention; time pressure reduced checking. Teachers must therefore practice (and teach) verification as behavior under constraints, not as an attitude alone. 44	Strong for behavioral trainability; implies need for teacher modeling under realistic constraints.

Synthesis: necessary and sufficient conditions for cultivating each human OS competency

A structured answer to A vs B

The evidence supports a **hybrid conclusion**:

- **A (conceptual understanding + activity design)** is *necessary*—teachers must know what they are aiming to cultivate and how to structure practice.
- **B (embodied practice)** is often a *minimum condition* for authentic cultivation when the competency involves (i) real-time judgment, (ii) transfer beyond school tasks, and/or (iii) moral legitimacy.

This follows from: (a) modeling mechanisms in cognitive apprenticeship, (b) transfer research emphasizing metacognition and dispositions, (c) empirical evidence that AI assistance can reduce agency and undermine independent performance, and (d) credibility/hidden-curriculum pathways through which students infer what “counts.” 53

Mechanism visualization: modeling → internalization → transfer in AI conditions

```
flowchart TD
T[Teacher practiced competence in AI-mediated tasks] --> M[Modeling: think-aloud + visible verification + ethical justification]
M --> S[Scaffolds: prompts, checklists, rubrics, worked examples]
S --> P[Student practice with feedback]
P --> R[Reflection: error analysis + postmortems + metacognitive logs]
R --> I[Internalization: habits/dispositions + monitoring routines]
I --> X[Transfer under AI presence: students self-initiate checks + act ethically]
AI[AI system: fluent outputs + occasional errors] -->|temptation|
OF[Cognitive offloading risk]
OF -->|without modeling| FAIL[Overtrust + shallow learning]
OF -->|with modeling+scaffolds| X
```

Competency-to-embodiment mapping table

Assumption note: no educational level was specified; the mapping targets **K-16**, with contextual caveats because much empirical GenAI research is currently concentrated in secondary and higher education and in specific national settings. 54

Human OS competency	What successful cultivation requires in AI-saturated contexts	Minimal embodiment required from teachers	Why embodiment is (or is not) necessary	PD implications (what to train and credential)
Question formulation / inquiry discipline	Students can (a) pose investigable questions, (b) iterate prompts as inquiry (constraints, assumptions, tests), (c) maintain epistemic curiosity rather than answer-seeking. ⁵⁵	Moderate: teachers must routinely demonstrate question refinement and show how prompts encode hypotheses and constraints. ⁵⁶	Inquiry is learned as a <i>practice</i> ; prompt engineering is explicitly described as iterative interaction requiring distinct skills because LLMs can hallucinate convincingly. Students need to see iteration in action. ⁵⁷	PD should include: prompt-to-question translation drills; teacher “live inquiry demos”; assessment rubrics for question quality and constraint articulation. ⁵⁸
Critical evaluation of AI outputs	Students can (a) verify claims, (b) triangulate sources, (c) detect hallucinations/bias, (d) resist automation bias under time pressure. ⁵⁹	High: teachers must personally practice verification/ lateral reading and show error detection routinely. ⁶⁰	Verification is behavioral and contextual; inoculation can shift behavior, but time pressure reduces checking; teacher modeling is needed to normalize verification as default action. ⁶¹	PD should credential: “verification fluency” (lateral reading routines, source evaluation), classroom protocols for AI-output checking, and assessment tasks requiring evidence trails. ⁶²

Human OS competency	What successful cultivation requires in AI-saturated contexts	Minimal embodiment required from teachers	Why embodiment is (or is not) necessary	PD implications (what to train and credential)
Ethical judgment and responsible use	Students can reason about authorship, integrity, bias, privacy, and responsibility; can justify choices, not just comply with rules. ⁶³	High: teachers must model ethical deliberation, transparency, attribution practices, and boundary setting in their own AI use. ⁶⁴	Moral norms are transmitted through hidden curriculum and credibility; rules without modeled practice risk hypocrisy and norm erosion. Character/ SEL syntheses show benefits of school-based programs, but teacher enactment is central to implementation credibility. ⁶⁵	PD should include: dilemma-based rehearsals (grading, feedback, authorship), transparent disclosure norms, and community moderation of edge cases; align with institutional policy but prioritize reasoning. ⁶⁶
Adaptive learning under technological uncertainty	Students develop SRL/meta-learning: planning, monitoring, strategy updating, epistemic humility, and resilience when tools change. ⁶⁷	Moderate-high: teachers should demonstrate reflective updating and admit uncertainty while showing learning strategies. ²¹ ⁶⁸	SRL and metacognitive monitoring are strongly implicated in transfer; AI accelerates change and increases uncertainty, making "reflection-in-action" a core modeled behavior.	PD should train: reflective practice routines, evaluation of AI tools over time, and collaborative norm-setting (since subjective norms may be weak). ⁶⁹
Keeping the human above the machine	Students maintain agency, creativity, and responsibility; AI is instrument, not authority; humans remain accountable for decisions and meaning. ⁷⁰	High: teachers must enact agency (human-in-the-loop judgment) and explicitly resist offloading of core thinking and moral responsibility. ⁷¹ ⁷²	RCT evidence suggests learners rely on AI assistance and don't fully replace it with SRL substitutes; without teacher-modeled agency, the machine becomes the default executive function.	PD should focus on: designing "AI-present but agency-required" tasks, assessment redesign, and teacher role formation as epistemic mentors/cognitive architects. ⁷³

Counterarguments and when facilitation without embodiment may suffice

Counterargument: “Teachers can facilitate competencies they do not personally practice; good curriculum is enough.”

There is partial truth here: well-designed scaffolds, worked examples, and guided practice can produce meaningful learning gains even when the teacher is not the domain’s top performer, especially for structured or near-transfer outcomes. ⁷⁴

However, the evidence reviewed suggests three boundary conditions where “design-only” facilitation becomes brittle:

1. **When the competency is dispositional and requires self-initiation** (e.g., verifying AI outputs without prompting), activity completion can occur without habit formation. ⁷⁵
2. **When AI introduces powerful incentives for offloading**, learners may optimize for output quality rather than cognitive growth unless adults normalize the growth-oriented stance by enacting it. ⁷⁶
3. **When moral legitimacy matters**, student perception of hypocrisy or avoidance can degrade the normative authority needed for ethical cultivation; credibility research and hidden curriculum theory make this risk non-trivial. ⁶⁴

Counterargument: “Requiring teacher embodiment is unfair and may create gatekeeping.”

This is a serious concern. “Embodiment” should not be operationalized as a moral purity test or as constant tool adoption. A defensible standard is **minimum viable embodiment**: enough lived practice to demonstrate key epistemic moves (checking, revising, attributing, setting boundaries), not maximal technical fluency. ⁷⁷

Policy and professional development implications

A practical implication is that teacher development should shift from “tool training” to **epistemic practice studios**: contexts where teachers repeatedly rehearse inquiry, verification, ethical deliberation, and agency under realistic AI conditions (time pressure, ambiguous tasks, imperfect outputs). The rationale is sharpened by evidence that teacher readiness includes concerns about autonomy, social-emotional learning, and role threats, not only technical integration skills. ⁶⁹

A 2026 competency framework for higher education explicitly distinguishes teachers’ own proficiency from their capacity to foster students’ critical engagement with GenAI and structures progression from basic to advanced levels—suggesting a model for credentialing teacher growth without demanding perfection. ⁷⁸

Credentialing could therefore focus on observable performances, for example:

- demonstrating an AI-verification routine (with evidence trails),
- facilitating an ethics dilemma with transparent reasoning,
- designing an “AI-present, agency-required” assessment,
- documenting reflective iterations of prompts and task designs across a term. ⁷⁹

Limitations and open research questions

The strongest limitation in the current evidence base is that **direct causal studies linking teacher embodied practice to student long-term human OS competencies in GenAI classrooms remain**

sparse, and much of the best-known causal evidence about learning harm/benefit comes either from non-teacher contexts (knowledge work surveys) or from preprints and early-stage field studies. ⁸⁰

Key open questions include:

- What dosage and form of teacher modeling is sufficient to produce durable verification habits under AI presence, especially in younger learners? ⁶¹
 - How should assessment systems measure “critical engagement with AI” without incentivizing performative compliance? ⁸¹
 - Which pedagogical “guardrails” reliably prevent negative transfer (improved performance now, degraded independence later), and how do teacher practices interact with those guardrails? ⁸²
 - How do teacher credibility and trust dynamics evolve when teachers and students both use AI, and what transparency norms protect legitimacy? ⁸³
-

Key sources most central to this report’s conclusions

- Cognitive apprenticeship and making thinking visible: ⁸⁴
 - Transfer and metacognitive monitoring: ⁸⁵
 - AI assistance and student agency (RCT): ³⁹
 - AI dependence and critical thinking (empirical): ⁴⁰
 - Critical thinking and GenAI (survey of knowledge workers): ⁴¹
 - Inoculation training and verification behaviors: ⁴⁴
 - Teacher readiness and PD gaps: ⁶⁹
 - Teacher credibility and legitimacy: ⁸⁶
 - Epistemic authority disruption framing: ³³
-

1 4 33 70 Epistemic authority and generative AI in learning spaces: rethinking knowledge in the algorithmic age

<https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2025.1647687/pdf>

2 23 Cognitive apprenticeship and instructional technology /

https://www.ideals.illinois.edu/items/18005/bitstreams/64487/data.pdf?utm_source=chatgpt.com

3 5 8 13 16 38 53 55 84 *

https://files.eric.ed.gov/fulltext/ED338729.pdf?utm_source=chatgpt.com

6 22 28 67 68 75 85 Teaching critical thinking for transfer across domains. ...

https://pubmed.ncbi.nlm.nih.gov/9572008/?utm_source=chatgpt.com

7 12 19 32 39 71 72 76 82 <https://www.sciencedirect.com/science/article/pii/S0360131523002440>

<https://www.sciencedirect.com/science/article/pii/S0360131523002440>

9 44 61 79 <https://www.sciencedirect.com/science/article/pii/S2666557326000066>

<https://www.sciencedirect.com/science/article/pii/S2666557326000066>

10 14 36 64 65 83 86 <https://www.tandfonline.com/doi/abs/10.1080/03634520903131154>

<https://www.tandfonline.com/doi/abs/10.1080/03634520903131154>

11 47 69 Teacher readiness for generative AI: A Theory of Planned Behaviour approach - ScienceDirect
<https://www.sciencedirect.com/science/article/pii/S2590291125010824>

- 15 21 <https://studenthublive.open.ac.uk/sites/studenthublive.open.ac.uk/files/reflection%20practitioner%20-%20schon.pdf>
<https://studenthublive.open.ac.uk/sites/studenthublive.open.ac.uk/files/reflection%20practitioner%20-%20schon.pdf>
- 17 18 https://archive.org/download/experienceeducat00dewe_0/experienceeducat00dewe_0.pdf
https://archive.org/download/experienceeducat00dewe_0/experienceeducat00dewe_0.pdf
- 20 <https://fsi-ebcao.princeton.edu/document/349>
<https://fsi-ebcao.princeton.edu/document/349>
- 24 <https://pmc.ncbi.nlm.nih.gov/articles/PMC9733593/>
<https://pmc.ncbi.nlm.nih.gov/articles/PMC9733593/>
- 25 https://physics.uwyo.edu/~rudim/S14_JEduPsych_DoesDiscoveryLrngEnhance.pdf
https://physics.uwyo.edu/~rudim/S14_JEduPsych_DoesDiscoveryLrngEnhance.pdf
- 26 **Experimental and Quasi-Experimental Studies of Inquiry ...**
https://www.semanticscholar.org/paper/Experimental-and-Quasi-Experimental-Studies-of-Furtak-Seidel/cac0e7181d70bf75c4f39262cc28dc5521a55470?utm_source=chatgpt.com
- 27 https://knilt.arcc.albany.edu/images/9/9b/Critical_thinking_.pdf
https://knilt.arcc.albany.edu/images/9/9b/Critical_thinking_.pdf
- 29 74 https://www.danamillercotto.com/uploads/4/7/7/2/47725475/barbieri_et_al_2023_we_meta-analysis.pdf
https://www.danamillercotto.com/uploads/4/7/7/2/47725475/barbieri_et_al_2023_we_meta-analysis.pdf
- 30 https://arc.educationapps.vic.gov.au/bd9b9586-fdc0-43af-aee7-eae87feb3fdc/wittwer_renkl_epr_2010.pdf.rsf
https://arc.educationapps.vic.gov.au/bd9b9586-fdc0-43af-aee7-eae87feb3fdc/wittwer_renkl_epr_2010.pdf.rsf
- 31 **Teaching for Transfer**
https://files.ascd.org/staticfiles/ascd/pdf/journals/ed_lead/el_198809_perkins.pdf?utm_source=chatgpt.com
- 34 46 <https://arxiv.org/abs/2311.05232>
<https://arxiv.org/abs/2311.05232>
- 35 <https://www.tandfonline.com/doi/full/10.1080/21532974.2023.2247480>
<https://www.tandfonline.com/doi/full/10.1080/21532974.2023.2247480>
- 37 <https://www.mdpi.com/2227-7102/14/5/477>
<https://www.mdpi.com/2227-7102/14/5/477>
- 40 <https://www.sciencedirect.com/science/article/pii/S0001691825010388>
<https://www.sciencedirect.com/science/article/pii/S0001691825010388>
- 41 80 https://www.microsoft.com/en-us/research/wp-content/uploads/2025/01/lee_2025_ai_critical_thinking_survey.pdf
https://www.microsoft.com/en-us/research/wp-content/uploads/2025/01/lee_2025_ai_critical_thinking_survey.pdf
- 42 <https://papers.ssrn.com/sol3/Delivery.cfm/4895486.pdf?abstractid=4895486&mirid=1>
<https://papers.ssrn.com/sol3/Delivery.cfm/4895486.pdf?abstractid=4895486&mirid=1>
- 43 <https://pmc.ncbi.nlm.nih.gov/articles/PMC12255134/>
<https://pmc.ncbi.nlm.nih.gov/articles/PMC12255134/>
- 45 **Reading Less and Learning More When Evaluating Digital ...**
https://journals.sagepub.com/doi/10.1177/016146811912101102?utm_source=chatgpt.com

- ⁴⁸ ⁵⁴ ⁷⁷ Artificial intelligence in teaching and teacher professional development: A systematic review - ScienceDirect
<https://www.sciencedirect.com/science/article/pii/S2666920X24001589>
- ⁴⁹ <https://d-nb.info/1353113094/34>
<https://d-nb.info/1353113094/34>
- ⁵⁰ <https://www.sciencedirect.com/science/article/abs/pii/S1096751625000314>
<https://www.sciencedirect.com/science/article/abs/pii/S1096751625000314>
- ⁵¹ <https://link.springer.com/article/10.1007/s11528-025-01123-8>
<https://link.springer.com/article/10.1007/s11528-025-01123-8>
- ⁵² <https://www.sciencedirect.com/science/article/pii/S0742051X25001659>
<https://www.sciencedirect.com/science/article/pii/S0742051X25001659>
- ⁵⁶ ⁵⁷ ⁵⁸ Prompt engineering as a new 21st century skill
<https://openscience.ub.uni-mainz.de/server/api/core/bitstreams/f52e7f38-3e12-49e0-85c3-7dcf4fc4542b/content>
- ⁵⁹ ⁶⁰ ⁶² <https://www.sciencedirect.com/science/article/abs/pii/S2352250X23001823>
<https://www.sciencedirect.com/science/article/abs/pii/S2352250X23001823>
- ⁶³ <https://www.tandfonline.com/doi/full/10.1080/03057240.2022.2060196>
<https://www.tandfonline.com/doi/full/10.1080/03057240.2022.2060196>
- ⁶⁶ <https://www.mdpi.com/3042-8130/1/1/2>
<https://www.mdpi.com/3042-8130/1/1/2>
- ⁷³ ⁷⁸ ⁸¹ Empowering university teachers in higher education: A generative AI-responsive competency framework - ScienceDirect
<https://www.sciencedirect.com/science/article/pii/S2666920X26000032>