INSTITUTE FOR FUTURE INTELLIGENCE | AI IN EDUCATION SERIES
blog.ai-steps.org • April 2026

The Effort Paradox:

Why How Students Use AI Matters More Than Whether They Use It

By Dylan A. Bulseco, Ph.D.

Two independent randomized controlled trials now show that AI assistance can reduce knowledge retention by the equivalent of a full letter grade — and erode a student's willingness to persist through a hard problem in as little as fifteen minutes. But the same research also reveals an encouraging nuance: it is not AI access that causes harm. It is a specific pattern of use. This post explains what the science shows and what educators can do about it.

1. Two Studies Every Educator Should Know

The debate over AI in classrooms has produced a lot of opinion and not much rigorous evidence. Two recently published randomized controlled trials (RCTs), which is considered to be the gold standard of causal research, has changed that. While they were conducted independently, with different populations and tasks, their findings point in the same direction.

Study 1: Does AI hurt long-term retention?

Barcaui (2025) randomly assigned 120 undergraduates to either use ChatGPT as a study aid or rely entirely on traditional resources such as textbooks, library databases, and standard web search in order to prepare a presentation on AI and machine learning concepts over a period of two weeks. Neither group knew a surprise test was coming.

Forty-five days later, the students were tested on that material. Students who studied the traditional way scored 6.85 out of 10 on average while students who used ChatGPT scored 5.75 — an eleven-percentage-point gap which, under most grading scales, constitutes a full letter grade. The effect size (Cohen's d = 0.68) is considered medium-to-large in educational research. Roughly three-quarters of traditional learners outscored the average AI-assisted learner.

One detail stands out: the AI group spent 45% less time studying (3.2 hours vs. 5.8 hours). But when the researchers statistically controlled for study time, the AI group still underperformed. Less time was part of the story but the **quality of cognitive engagement was the bigger issue**.

Study 2: Does AI erode persistence?

Liu et al. (2026) ran three large-scale experiments with a combined sample of 1,222 participants who worked on fraction arithmetic or on SAT reading comprehension. Half had access to an AI assistant during a brief practice session and the other half did not. Then the AI was removed, and everyone worked on identical test problems alone.

The results were striking. AI-assisted participants solved significantly fewer problems correctly — a consistent finding across all three experiments and both content domains. But the persistence finding may be more telling: participants were allowed to skip any problem with no penalty. In the reading comprehension study, 8% of AI-assisted participants skipped test problems, compared to just 1% in the control group which is an eight-fold difference. **After working with AI, students were less willing to struggle**.

The most actionable finding from this study came from a closer look at how students actually used the AI. Those who sought direct answers showed the largest deficits. Those who used AI for hints or clarifications showed substantially smaller decrements. Students who had AI access but chose not to use it performed comparably to the control group. **Access was not the problem. The behavioral pattern was**.

> *Barcaui (2025): Students know less — AI-assisted learners retained 11% less knowledge at 45 days. Liu et al. (2026): Students try less — AI-assisted learners were 8x more likely to skip difficult problems.*

2. Why This Happens: Three Mechanisms Worth Understanding

These findings are not surprising to researchers who study human memory and learning. Three well-established frameworks all developed prior to transformer generative AI models converge on the same prediction: **tools that remove effortful thinking from the learning process also remove the mechanism that makes learning durable**.

Cognitive Offloading

We naturally delegate cognitive work to external resources when those resources are available. That is usually adaptive and is why we write things down, use calendars and calculators. These tools help carry out specific tasks. The problem arises when the work we offload is also the work that drives learning. Retrieval, synthesis, and working through a problem are effortful precisely because they build strong memory traces and this is precisely what LLMs are good at. When AI handles those operations, the encoding remains shallow.

Researchers have called a related dynamic the "Google effect": people who believe information is retrievable online are less likely to encode it in memory. Generative AI takes this further. It does not just retrieve facts it synthesizes, explains, and reasons through material. There is very little for the learner's cognitive system to do.

Desirable Difficulties

Memory researcher Robert Bjork coined the term "desirable difficulties" to describe learning conditions that feel harder in the moment but produce stronger long-term retention. These include things like spacing practice out over time, mixing different problem types, retrieving information from memory rather than re-reading it. The difficulty is desirable because it engages the processes that build durable knowledge.

Related research by Kapur (2008) on "productive failure" found that students who struggle with a problem before receiving instruction actually learn more than students who receive instruction first even when the initial attempts are wrong. The struggle itself is part of the mechanism. **AI that delivers the answer before that struggle eliminates the learning value of the attempt.**

Borrowed Competence and Shifting Effort Tolerance

Barcaui (2025) introduces a useful concept he calls "borrowed competence": when a student reads an AI's articulate, well-organized output, they experience a subjective sense of understanding even if the understanding belongs to the model and is not due to their own accomplishments. The student's feeling of mastery does not reflect their own knowledge. This creates a metacognitive blind spot: students finish a study session feeling confident without having built the cognitive architecture that would support retention.

Liu et al. (2026) add a related mechanism for the persistence findings: when AI routinely resolves problems in seconds, students recalibrate their sense of how much effort normal cognitive work requires. Subsequent unaided work then feels abnormally hard by comparison and skipping it becomes more acceptable. Each instance of AI assistance shifts this baseline a little further.

3. What Educators Can Do: Five Evidence-Based Strategies

The critical insight from Liu et al. (2026) is that harm was concentrated in a specific behavioral pattern, seeking direct answers, not in AI access per se. That means **the goal is not prohibition; it is structured use**. The following **five strategies** are grounded in the same cognitive science that explains the problem.

1

Struggle First, AI Second

Require students to engage with material independently before AI becomes available. Write a paragraph in your own words, attempt a problem, or identify a specific gap in your understanding and only then, open the chat. This preserves the effortful encoding that drives retention while still allowing AI to clarify, compare, and extend. This is similar to a student asking “what is on the test” vs. seeking to understand concepts that will help them prepare for a test.

2 **Prescribe Hints, Not Answers**

Guide students to ask AI for the next step or a useful analogy rather than a complete explanation. Provide prompt templates: "What concept am I likely missing?" or "What question should I be asking myself here?" This keeps the student doing the cognitive work while AI plays a scaffolding role which is exactly the pattern of behavior the research shows is far less harmful.

3 **Flip the Dynamic: Students as Critics**

Ask students to produce a draft or solution first, then have them prompt AI for an alternative or critique then evaluate that AI output for errors, omissions, and limitations. This requires genuine domain knowledge to execute well and addresses the borrowed competence problem directly: students must know enough to identify where the AI is wrong.

4 **Build Retrieval Practice Into Your Routine**

Begin class sessions with brief low-stakes questions on material from two or three weeks prior — not the most recent lesson. This spacing is what strengthens memory traces over time. AI can support this by generating varied practice questions after students have done initial independent study, leveraging its strengths without bypassing the effort that matters. Furthermore it reinforces the scaffolding and helps strengthen connections between topics.

5 **Assess Process, Not Just Product**

A polished essay or correct problem set no longer reliably signals that a student has learned. Oral examinations, in-class demonstrations, supervised written responses, and reflective journals documenting student reasoning all reveal what students actually know. Process-focused assessment is more time-intensive and in the era of LLMs, definitely more valid.

4. A Note on What This Research Does — and Does Not — Show

These studies document the harm of a specific pattern of AI use: unrestricted, answer-seeking interaction during initial learning. They are not a case against AI in education generally.

Other research has found that structured, scaffolded AI integration can support differentiated instruction, increase engagement, and help students who struggle with foundational concepts. The distinction the evidence draws is not AI versus no AI, but rather AI that replaces cognitive effort versus AI that supplements it after students have done meaningful work.

A calculator does not impair mathematical reasoning if students already understand the operations it performs. The emerging consensus on AI in education suggests analogous logic applies: sequence, structure, and pedagogical intent determine the outcome.

> *The core principles of effective learning — effortful retrieval, generative responding, productive struggle, spaced practice — are not rendered obsolete by generative AI. The current evidence suggests they are more important to protect than ever.*

Conclusion

The evidence from these two randomized controlled trials is clear enough to act on. When students use AI as a direct answer source during initial learning, they remember less, try less, and have limited awareness that either is happening. The mechanisms are well understood, the effect sizes are meaningful, and the findings replicate across different task domains and student populations.

For educators in community colleges and K-12 settings, the practical path forward runs through the same cognitive science that has always guided good instruction: make students think first, build in retrieval, make practice genuinely difficult, and assess in ways that reveal what students actually know. AI can support each of those goals if it is introduced at the right point in the learning sequence.

The question is not whether students will use AI. They will. The more productive question and one educators have the expertise to answer is: **how do I design my courses so that AI use strengthens my students' thinking rather than replacing it?**

References

- Barcaui, A. (2025). ChatGPT as a cognitive crutch: Evidence from a randomized controlled trial on knowledge retention. Social Sciences & Humanities Open, 12, 102287. https://doi.org/10.1016/j.ssaho.2025.102287
- Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher et al. (Eds.), Psychology and the real world (pp. 56-64). Worth Publishers.
- Kapur, M. (2008). Productive failure. Cognition and Instruction, 26(3), 379-424.
- Liu, Y., et al. (2026). AI assistance and its effects on student persistence and independent performance: Three randomized controlled experiments. [Preprint / in 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
- Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. Science, 333(6043), 776-778.