

Two years of evidence:

SchoolAI makes students think

schoolai



The bottom line:

Over two years, students using SchoolAI demonstrated 28% greater critical thinking. Higher-level reasoning more than doubled. The gains held across every subject and grade level studied.

+28%
boost in critical thinking

Why this matters

Critical thinking has always been hard to teach at scale. Studies going back decades show that most classroom questions focus on memorization and recall, not because teachers don't value deeper thinking, but because facilitating higher-order thinking across an entire class is genuinely difficult. We default to what scales.

AI promised to change that, precisely because it can operate at scale. But most [general-purpose AI](#) tools were built for productivity, not pedagogy. They are designed to produce answers, and the concern is legitimate: when AI does the reasoning, students don't have to.

SchoolAI was built on a different premise, specifically designed for education. Rather than optimizing for fast answers, it was purpose-built to deepen thinking.

This study set out to ask whether SchoolAI's approach can do the thing that has never scaled: coach every student toward analysis, evaluation, and genuine understanding.

The results suggest it can.

What we did

We analyzed 13,882 student-AI conversations from 82 teachers in the Jordan school district in Utah (55,000 students, 68 schools) at two points in time: October 2023, shortly after SchoolAI launched, and October 2025, after two years of use. By tracking the same 82 teachers across both timepoints, we could measure how student engagement with AI changed as teachers gained experience with the platform.

Each conversation was scored on a four-level critical thinking scale based on Bloom's Taxonomy, from basic recall to analysis, evaluation, and creation.

You've been studying the water cycle and how its parts connect to life on Earth. If evaporation stopped completely, what would happen to life on Earth over time?



All the plants and animals would die because there would be no more rain and everything would dry up!

Level

Definition

Example Indicators

1	No observable critical thinking	"What is...?", "ok", choice selections ("A", "2"), off-topic conversation, pasted content without engagement
2	Remember + Understand	"The answer is...", "First I..., then I...", restating information, seeking clarification
3	Apply + Analyze	"I multiplied because...", "This works when...", comparing approaches, explaining reasoning
4	Evaluate + Create	"Method A is better because...", "I noticed a pattern...", revising work, synthesizing ideas

What we found

Critical thinking improved dramatically.

Level	Oct 2023 % (n)	Oct 2025 % (n)	Change
1	76.7% (7,280)	58.0% (2,694)	-18.7%
2	11.8% (1,147)	16.5% (663)	+4.6%
3	8.2% (557)	16.0% (905)	+7.8%
4	3.2% (207)	9.5% (429)	+6.3%

Mean critical thinking scores rose 28.3% from October 2023 (M = 1.38, SD = 0.49) to October 2025 (M = 1.77, SD = 0.69). Levels 3 and 4, the deepest levels of analytical and evaluative thinking, more than doubled.

These gains weren't a fluke of who was using SchoolAI. Statistical models confirmed the improvements held after accounting for differences in subject area and grade level. Teachers who actively designed learning experiences on the platform, creating six or more unique spaces, showed significant gains, while those with minimal usage did not, confirming that improvement was tied to meaningful engagement with SchoolAI rather than external factors. In 2023, usage in SchoolAI mirrored industry patterns, lower-order thinking. By 2025, that had flipped: nearly half of all conversations showed critical thinking.

Not all "Level 1" is the same

Level 1 doesn't always mean a student wasn't thinking. Many Level 1 conversations reflect intentional lesson designs, such as book selection tools, historical character simulations, or branching scenarios where students make choices rather than explain reasoning. The critical thinking still happened, but it's not visible in the text. By 2025, choice-based activities made up 36.2% of Level 1 conversations, up from 20.1%, suggesting that even within the lowest level, the nature of engagement was shifting. The remaining Level 1 conversations in 2025 increasingly reflect pedagogically intentional activities rather than passive or disengaged behavior.

One Piece of the Puzzle

This study documents what changed over two years, but not why. Critical thinking gains reflect a complex interplay of factors, including how teachers design learning experiences, how the AI responds, and how students engage over time. Our companion report, [Design Matters: How SchoolAI Architecture Shapes Student Thinking](#), examines exactly this. Understanding these factors is essential to supporting and sustaining student outcomes.

The takeaway

In two years, students went from a classroom pattern dominated by recall to one where nearly half of all AI conversations showed higher-level thinking. That didn't happen because students had access to AI. It happened because SchoolAI gave teachers the tools to design learning experiences where AI guides the thinking instead of replacing it. That's the difference between AI in classrooms and AI that actually improves learning.

This quick report is a shortened version of the [full report](#) that was externally reviewed by [Instructure](#) and received ESSA Tier 3 Certification.

The evidence is clear: thoughtfully designed AI, like SchoolAI, can raise the cognitive bar for students at scale.

