

Teacher-Guided AI Moves Ahead as Schools Rebuild Lessons, Tutoring, and Assessment

AI in EdTech Weekly

2026-06-15

Teacher-Guided AI Moves Ahead as Schools Rebuild Lessons, Tutoring, and Assessment

By AI in EdTech Weekly • June 15, 2026

This week’s strongest signal is that AI works best in education when it is structured, teacher-guided, and tied to clear learning tasks. New school-built tools, district deployments, and policy test beds show where adoption is accelerating—and where feedback, assessment, cost, and governance still lag.

The clearest signal this week: structured AI is outperforming generic AI use

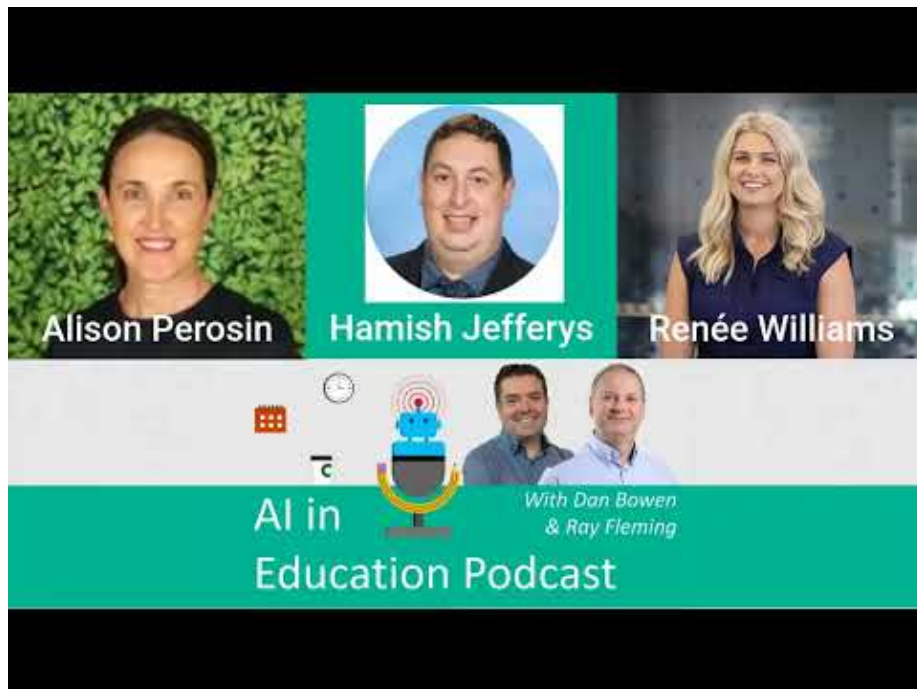
In Sierra Leone, Google DeepMind positioned AI as a response to teacher shortages, describing it as a partner that can extend educators’ reach without replacing them [1]. Over eight weeks, students increasingly used Gemini to understand concepts rather than just get answers, with problem-solving queries rising from 68% to 90% [2]. EdSurge also pointed to a Sierra Leone study in which a one-day AI training for secondary teachers was followed by math gains equivalent to more than a year of additional schooling [3].

“AI can act as a partner to support educators in these environments – amplifying their reach without replacing their essential expertise and skills.” [1]

The contrast with generic chatbot use is getting sharper. Standard free LLMs can lower brain activity and retained learning by encouraging what one report called “cognitive surrender,” while a carefully designed AI tutor in an undergraduate physics course produced twice the learning gains of active, in-person instruction [3]. Estonia’s emerging policy follows the same logic: students build foundational knowledge first, then use AI later in the learning process for feedback and assisted learning; earlier grades are deliberately excluded for now [3].

Schools are building tighter AI workflows instead of relying on public chatbots

In Australia, some schools are already doing this themselves. In Broken Bay Diocese, a Year 6 student built a science agent inside a controlled “kids’ pool” environment; it checks a learner’s level, adapts explanations or tests, and can even add engagement cues like jokes. The class later adopted the agent because it worked across different learning needs [4].



Flourishing in the Age of AI: Schools Putting Humans First (1:53)

Another school built a secure Gemini + Apps Script tool that combines GPA, testing, and timetable data so teachers can query class or student breakdowns and get differentiation suggestions without moving student data into public systems [4]. At Cathedral College in Rockhampton, a lesson-starter agent was grounded in school teaching frameworks, curriculum documents, and teacher-contributed examples, but its creator emphasized that cultural preparation came first and that the tool was not meant to replace human mentoring [4].

“It can’t exist on its own. It’s not meant to be a standalone agent or replace a human coach or mentor.” [4]

District tutoring programs are moving in the same direction. Newark Public Schools received \$400,000 from New Jersey to expand high-impact tutoring that uses AI with teacher oversight for math and reading [5]. District leaders said they expanded Khanmigo after pilot users showed math-score improvement [5].

The next wave of tools is more lesson-native than chatbot-native

Microsoft’s new Learning Zone turns educator prompts, uploaded files, or vetted resources such as OpenStax into interactive lessons in minutes [6]. The lessons combine bite-sized content slides with multiple exercise types, immediate feedback, retries, and conditional “nested” slides that give students extra practice when they miss a concept [6]. Teachers can use it for topic introductions, wrap-ups, flipped learning, or live instruction with anonymous aggregated knowledge checks, then assign lessons through codes, links, Teams, or an LTI-compatible LMS and review performance reports afterward [6].



Microsoft Learning Zone - Hands on webinar for Teachers (June 2026) (45:07)

The capability is notable, but the limits matter too. Students can access lessons in a browser on any device, while lesson generation currently requires a Copilot+ PC, with a broader trial planned [6]. Microsoft also says more generation languages and an in-class teaching mode are coming [6].

Assessment is still where AI most clearly hits its limits

The week’s most useful assessment research separates scoring from feedback. In a randomized trial across 178 schools in Brazil, AI essay scoring performed at the level of human review; students improved by about a tenth of a standard deviation whether essays were scored by AI alone or AI plus human graders [7].

But feedback is a different task: it requires identifying what a specific student got wrong in the context of that student’s reasoning [7].

That gap shows up across multiple studies. In middle-school math, the model that scored best produced teacher-preferred feedback only 12% of the time, while GPT-4 produced the most trusted feedback and the worst scores [7]. On science work, LLMs matched teachers on next-step “Feed Forward” guidance but lagged on “Feed Back” that diagnoses the student’s specific reasoning error, scoring 3.05 versus teachers’ 3.52 out of 5 [7]. In a physics-feedback study, about 20% of AI responses were inaccurate, and students rated wrong feedback as just as accurate as correct feedback [7].

This helps explain why AI is reallocating teacher time rather than eliminating the teacher role. In Brazil, AI scoring gave teachers about 30% more one-on-one writing conferences without increasing workload [7]. And it helps explain Justin Reich’s warning that AI is most useful when the user already has enough domain knowledge to separate strong output from confident nonsense [8].

Real classrooms are already adapting. Reich’s 120-interview project found widespread homework bypass, with some students deciding which assignments are important enough to do themselves and teachers responding with everything from rewrites and detectors to harder AI-required tasks [8]. Teachers on Reddit describe moving essays and tests in-class or on-demand because “anything that goes home is AI’d now” [9]. Another teacher working with younger students reported near-identical AI-generated responses on a research assignment [10].

Higher ed and policy are moving from experimentation to governance

In higher education, roughly three-quarters of faculty say students use GenAI to write essays and papers, and roughly the same share of faculty use GenAI themselves [11]. But most institutions are still in a wait-and-see phase or running patchwork experiments, not seeing broad gains in learning or efficiency [11]. That is pushing redesign in two directions: tougher assessments such as live oral defenses, presentations, and more rigorous feedback loops [11], and better student-support systems that connect academic, financial, and well-being data while reducing administrative overhead [11]. It also aligns with the argument that higher education and workforce systems need continuous reskilling, more personalized learning, and more real-world experience as AI changes work [12, 11].

Governance is becoming more explicit. EDUCAUSE described low-risk AI uses such as brainstorming, tutoring, translation, summarization, and simple coding; medium-risk uses such as course design, grading, student feedback, and administrative tasks; and high-risk uses involving student records, HR data, or financial information [13]. For community colleges, the warning is that commercial tools may encode assumptions that do not fit part-time, working, or

caregiving students, which can automate weak judgments at scale [14]. Recommended safeguards include contextual auditing, vendor transparency, and collaborative governance with faculty and student advocates [14].

The same shift is happening at system level. England’s Department for Education has funded 16 edtech firms to build trustworthy AI tools for lesson planning and marking using a national curriculum data/content store prototype [15]. A £1 million pilot has now expanded into a £23 million, four-year program recruiting 1,000 schools and colleges as test beds for AI and edtech [15]. Google DeepMind and Edy have announced a randomized trial of Learn LM with 1,500 students across 10 schools in England [15], while OpenAI is testing a learning-outcomes measurement suite with 20,000 students in Estonia and Anthropic and CodePath are running a 15-month classroom study across thousands of students [15].

Scale, though, is not the same as settled evidence. Ben Williamson argues that these programs can privilege signals of impact and scalability over broader forms of evidence, while reshaping classrooms into measurable testing sites for experimental AI products [15]. That warning matters because AI economics are not like normal software: every use carries inference costs, private or local deployments add storage, cybersecurity, hardware, networking, and technical expertise, and districts still have few examples of what universal access would actually cost [16]. UNESCO’s Digital Transformation Collaborative is one example of the response, framing digital transformation around coordination, connectivity, cost, capacity, content, and data [17].

What This Means

- **For school systems:** Separate AI for tutoring and practice, AI for lesson creation, and AI for assessment. The same tool may score reliably and still fail at diagnostic feedback [7].
- **For instructional leaders:** Training and workflow design matter more than access alone. Sierra Leone’s gains followed teacher preparation, and Australian implementations started with secure environments and cultural work—not just tool rollout [3, 4].
- **For higher ed and L&D teams:** Access without redesign leads to patchwork. The durable opportunities are stronger assessment, better student support, continuous reskilling, and more real-world AI use through projects, internships, co-ops, and apprenticeships [11, 12].
- **For buyers and investors:** Ask what must be true locally—curriculum fit, teacher prep, infrastructure, language, inclusion, data protection, and affordability—before treating pilot results as scalable [17].
- **For learners and families:** In this week’s coverage, AI literacy was defined less as basic tool use and more as questioning outputs and evaluating reliability; foundational knowledge still determines whether AI helps or misleads [18, 8, 3].

Watch This Space

- **Teacher-made microtools:** “Vibe coding” is lowering the barrier for teachers to build task lists, translation workflows, dashboards, and practice games. One fourth-grade teacher reported an AI-built review game that led to students scoring five points higher on average with no retests [19].
 - **More frontier-lab trials inside education systems:** The DfE test-bed expansion, Learn LM trial, OpenAI’s Estonian measurement suite, and Anthropic’s CodePath study will shape both evidence and market expectations [15].
 - **AI-native school models:** Alpha World School’s launch points to a more ambitious version of AI-enabled schooling, pairing daily AI-driven academics with fieldwork in Kenya and Ecuador and research projects with university faculty [20].
 - **Agentic tools for self-directed learning:** NotebookLM’s new research companion can build a source repository from loose questions, surface its reasoning process, and export outputs in formats from charts to spreadsheets and documents [21, 22, 23].
 - **Cost and governance as adoption bottlenecks:** As pilots broaden, recurring inference costs, privacy demands, and local hosting decisions may matter as much as the model itself [16].
-

Sources

1. X post by @GoogleDeepMind
2. X post by @GoogleDeepMind
3. AI Won’t Replace Educators. But It is Changing How Students Learn.
4. Flourishing in the Age of AI: Schools Putting Humans First
5. New Jersey awards Newark \$400K to boost tutoring programs built on AI and high-impact sessions
6. Microsoft Learning Zone - Hands on webinar for Teachers (June 2026)
7. There’s More to AI Grading Than Scoring
8. AI in the Classroom — Why There Are No Best Practices Yet
9. r/Teachers comment by u/Delphgirl
10. r/Teachers post by u/Katpagla
11. Building an AI-Ready America: Higher Education in the Age of AI
12. Why Athletics are the ‘Shadow University’ and the Impact of AI on Degrees and Reskilling
13. How AI Is Changing Campus Cybersecurity: 4 Key Challenges | EDU-CAUSE Exchange
14. Invisible Infrastructure: Why AI Ethics Are the New Mandate for Community College Leaders
15. #EdTech26 Keynote 2 Dr Ben Williamson ‘Fabricating EdTech Futures’
16. Can Schools Afford an AI-First Future?

17. EdTech Partnerships Need Evidence-Informed Governance: Where do we start?
18. Education in the AI Era: Bridging the divide between school & the real world
19. Vibe Coding for Teachers: No Coding Skills Needed with Donnie Piercey
20. X post by @mackenzieprice
21. X post by @NotebookLM
22. X post by @NotebookLM
23. X post by @NotebookLM