A MAP OF THE RESEARCH ON AI AND LEARNING

How can AI enable *real learning?*

What the research actually says.

THE PROBLEM

AI can improve performance and *reduce learning*.

In a randomised study of about 1,000 high school maths students in Türkiye, students using a standard ChatGPT scored 48% higher than classmates without it while they had it. When access was taken away, they scored 17% lower than students who never had it. A 2025 paper in Nature Reviews Psychology makes the underlying point: performance gains are not the same as learning.

Better output is not the same as learning.

Source: Bastani et al. (2025), PNAS. On performance vs learning: Yan, Greiff, Lodge & Gašević (2025), Nature Reviews Psychology.

WHY IT HAPPENS

When AI does the thinking, *the learning does not happen*.

Most AI tools were built for work, not to optimise learning. At work, the goal is to finish the task with the least effort. But in learning, that effort is the point: it is what builds the capability. The task gets finished, but the understanding never develops. Researchers call this “metacognitive laziness”: the learner stops planning, monitoring and self-evaluating, because the AI always has an answer.

If the AI carries the effort, it also carries off the learning.

Source: Khosravi et al. (2026), Building AI Companions that Prioritise Learning over Performance; metacognitive laziness: Fan et al. (2025), BJET.

WHY EFFORT IS THE POINT

The struggle is not in the way of learning. *It is the learning.*

Learning scientists Robert and Elizabeth Bjork call these “desirable difficulties”: things like recalling an answer from memory, spacing your practice out, or trying a problem before you see the solution. They make learning feel harder now, but they make it last. When AI removes that effort, it can quietly remove the learning the effort produced. A 2026 review of 67 studies names the same mechanism, epistemic friction: without it, “AI-generated fluency can bypass the reflective struggle central to deep learning.”

Protect the difficulty that does the teaching.

Source: Bjork & Bjork, desirable difficulties; Li, Cui & Hagedorn (2026), Computers and Education: AI.

BUT DESIGN CHANGES EVERYTHING

The same technology can help or harm. *The design decides which.*

In a Harvard physics study, a purpose-built AI tutor designed with proper scaffolding beat in-person active learning by 0.73 to 1.3 standard deviations, two to three times the usual bar for a substantial effect in education research. In the Türkiye study, the standard ChatGPT left students worse off once it was removed, while a guardrailed tutor version avoided that loss entirely. Same models, opposite outcomes, depending on how they were designed.

The result is set by the design, not by the model.

Source: Kestin et al. (2025), Scientific Reports; Bastani et al. (2025), PNAS.

AND BY THE HUMAN BEHIND IT

Believing a human is paying attention changes *how hard we try*.

In a controlled study in a university creative-coding course, students received identical AI-generated feedback on their work. Those told it came from a human teaching assistant ran their code more, wrote more code, and spent more time on later work. They rated the feedback equally helpful either way. The effect on effort was large (d = 0.88 to 1.56). The content was the same.

Same words land differently when we believe a human wrote them.

Source: Morris & Maes (2026), Same Feedback, Different Source.

WHAT ONLY A HUMAN DOES

AI can help with the content. *It rarely touches the rest.*

Education does three things at once: it builds knowledge and skills (qualification), it helps you find your place among others (socialization), and it helps you become someone who thinks independently and takes responsibility (subjectification). AI tools mostly reach the first. They rarely, if ever, address the other two, and those are where a teacher does their deepest work.

AI can teach the content. A human helps you become someone.

Source: Gert Biesta, three functions of education; Wayne Holmes (2026), Learning to think in the AI era.

THREE WAYS AI CAN SHOW UP

An LLM, a tutor, and a learning companion are *not the same thing*.

An LLM answers your question. Faster work, less learning.

An AI tutor asks questions back, no matter what you actually need. Often frustration and drop-out.

An AI learning companion (Philippa Hardman calls it a “study mate”) remembers where you got stuck and pushes you towards the thinking you avoid. Capability that lasts.

Aim for a learning companion, not an answer machine.

Source: Khosravi et al. (2026), Building AI Companions that Prioritise Learning over Performance; “study mate” framing via Dr Philippa Hardman.

WHAT EDUCATION IS FOR NOW

AI does not shorten what there is to learn. *It lengthens it.*

Holmes (UNESCO) asks it directly: if generative AI is this powerful, do we still need to learn? His answer is yes, and the list grows: on top of what we wish to learn, we now need to learn AI's profound limitations, its impacts on human rights, social justice and the environment, and "perhaps most importantly, [to] learn how to think… critically." The World Economic Forum (WEF) keeps the balance: rote memorisation may matter less, but "the process of mastering knowledge continues to develop broader capabilities" such as grit, curiosity, communication and critical thinking, and assessment must evolve to capture them.

The emphasis moves from having answers to judging them.

Source: Wayne Holmes (2026), UNESCO Courier; WEF (2026), Assessments readiness signal.

The challenge is also systemic. *It runs through the tool, the classroom, and the system around them.*

The tool: the guardrailed tutor and the learning companion are instructional design built into software: scaffolding, answers withheld, help that fades as you grow. (Bastani et al., 2025; Khosravi et al., 2026)

The classroom: co-design tools with teachers rather than deploying them on teachers, and set tasks that make learners compare, justify and revise what the AI produces. In the 67-study review, that scaffolding is what separated gains from cognitive offloading. (OECD, 2026; Li, Cui & Hagedorn, 2026)

The system: AI only helps where the conditions are ready. As the WEF puts it, “learning outcomes will not be determined by technology itself, but by the conditions in which it is deployed”, and isolated fixes across policy, pedagogy and technology are unlikely to be sufficient. (WEF, 2026)

Conditions at each level shape outcomes at every other. Real progress needs all three moving.

Source: Bastani et al. (2025); Khosravi et al. (2026); OECD (2026), chapters 7 and 8; Li, Cui & Hagedorn (2026); WEF (2026), executive summary.

SO, HOW CAN AI ENABLE REAL LEARNING?

Protect the difficulty that does the teaching.

The result is set by the design, not by the model.

AI can teach the content. A human helps you become someone.

Aim for a learning companion, not an answer machine.

Isolated fixes are unlikely to be enough. The levels have to move together.

One test cuts through it all, the one Inara Scott's framework points to: ask who is doing the thinking, you or the AI. Keep yours alive, and use AI to learn, not just to finish.

Are you here for the output, *or to get better at what you do?*

Source: Inara Scott (2026), The AI Cognitive Pyramid, SSRN.

Where this comes from

A map of the research, drawn over a few months and compiled with the help of AI.

Bastani et al. (2025), Generative AI without guardrails can harm learning, PNAS
pnas.org/doi/10.1073/pnas.2422633122

Kestin et al. (2025), AI tutoring outperforms in-class active learning, Scientific Reports
nature.com/articles/s41598-025-97652-6

Yan, Greiff, Lodge & Gašević (2025), Distinguishing performance gains from learning when using generative AI, Nature Reviews Psychology
nature.com/articles/s44159-025-00467-5

Fan, Tang et al. (2025), Beware of metacognitive laziness, British Journal of Educational Technology 56(2), 489–530
doi.org/10.1111/bjet.13544

Robert & Elizabeth Bjork, desirable difficulties (learning-science framework)
bjorklab.psych.ucla.edu

OECD, Digital Education Outlook 2026 (teacher co-design: chapters 7 and 8)
oecd.org/en/publications/oecd-digital-education-outlook-2026_062a7394-en.html

World Economic Forum, Shaping the Future of Learning (2026; quotes from the executive summary and the Assessments section)
weforum.org/publications/shaping-the-future-of-learning-education-readiness-for-the-age-of-ai/

Khosravi et al. (2026), Building AI Companions that Prioritise Learning over Performance, arXiv
arxiv.org/abs/2605.04816

Morris & Maes (2026), Same Feedback, Different Source, arXiv
arxiv.org/abs/2602.11311

Li, Cui & Hagedorn (2026), The cognitive impact of ChatGPT in higher education (review of 67 studies), Computers and Education: AI
doi.org/10.1016/j.caeai.2026.100571

Inara Scott (2026), The AI Cognitive Pyramid, SSRN
papers.ssrn.com/sol3/papers.cfm?abstract_id=6353318

Gert Biesta, three functions of education (qualification, socialization, subjectification)
cited via Holmes

Wayne Holmes, Learning to think in the AI era, UNESCO Courier (April 2026)
courier.unesco.org/en/articles/learning-think-ai-era

Dr Philippa Hardman, From AI Tutors to AI Study Mates
drphilippahardman.substack.com/p/from-ai-tutors-to-ai-study-mates