1.2 Teaching About vs Teaching with AI

Take Home message

- ***AI is a satnav.*** AI can save the driver who already carries the map and prevent the one who doesn't from learning. The effect depends on the person, not the product.
- ***Teaching about AI is the precondition for Teaching with AI. Y***ou need to know the machine to check the answer the machine produced.
- **Knowing is not doing.** Teachers rate their AI knowledge above their ability to teach with it, and training alone does not close that gap.
- ***Teachers use AI backstage, not in class.*** Half of the teachers use AI regularly, under a third use it with students.
- ***Pedagogy beats technology.*** Transforming lessons beats chatting with AI; working in pairs beats working alone; feedback beats gamification.
- ***Better feedback is not better learning.*** AI comments judged clearly better than teachers' produced revisions that were no better at all. What decides is whether feedback is *taken up.*
- ***AI raises output, not self-directed learning.***
- ***Clear rules and heavy use go together.*** Teachers in schools with a written AI policy report saving more time, not less — the opposite of what most people expect.
- ***Guard your own map.*** *Relief becomes deskilling where AI absorbs the work that builds and maintains judgment.*

> *If your remember only one sentence from this page:* **AI is a satnav. It saves the driver who already has a map in her head and quietly stops the driver who hasn't from ever building one**

The satnav problem

To drive a London taxi, you must first pass a knowledge test: 320 set routes across the 113 square miles within six miles of Charing Cross, for which Transport for London allows up to two years before the examinations even begin.

Examined drivers also use a satnav, and nobody thinks less of them for it. The knowledge test did not become worthless when the device arrived. The satnav became the thing that makes the drive safe. A driver with the city in her head uses the satnav for what she cannot know: the accident on the bridge or tonight's traffic. And she notices when it routes her down a street that has been dug up since April.

Now put a newcomer in the same taxi with the same satnav. She arrives at every address. For a while, she is indistinguishable from the experienced driver. But following the satnav's instructions, she is building nothing. After a year of turning left when instructed, she still cannot cross the city without the device, and she cannot tell when it is wrong.

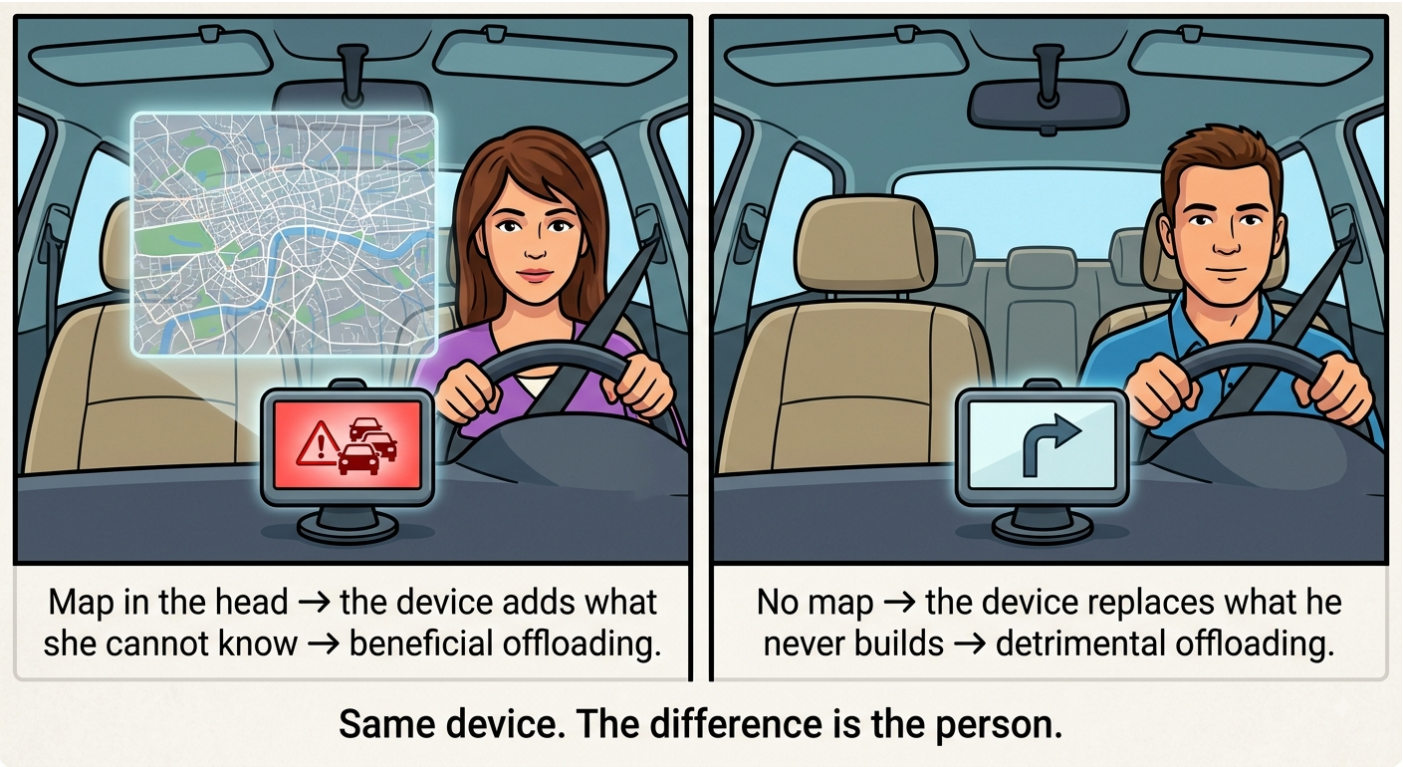

Fig. 1. The outcome depends on the driver. Image created with GoogleGemini.

Same satnav, two drivers, opposite outcomes. **The difference was never in the device.**

That is a decade of research on AI use in one image, and it holds without the image too. The effect of an AI tool depends far more on what the user already knows than on the tool itself. This is what researchers call a schema, the organized prior knowledge of a person. Where it is in place, handing work over to a machine frees you up. Researchers call this beneficial cognitive offloading. However, when knowledge needs to be developed, the same process can prevent it from forming. This is called detrimental cognitive offloading.

A study with around a thousand secondary students shows this happening in a mathematics classroom. One group practiced with an unrestricted chatbot, another without. While they had it, the chatbot group solved **48 % more practice problems!** Then came an exam with the chatbot taken away. Now, the same students who solved practice problems so well scored **17% worse** than the group that had never used the AI chatbot. A third group used an AI version built to withhold answers and ask questions instead. That group lost nothing.

That third group is the whole design question. A satnav that names the destination and helps you find the route teaches you the city. A satnav that says *turn left in 200 meters* does not, however good its data. Researchers call the difference guardrails, and the **design decides whether you help your students learn the city's map or simply find the destination.**

Please remember: **You have the knowledge. Your students, in the subject you teach, do not yet.** That asymmetry is the whole chapter.

Two floors, not two topics

In academic literature, these two areas are considered distinct fields. You could think of them as two floors of a building, with most teachers located on the ground floor and no staircase leading to the first floor.

When asked to rate their knowledge of AI, teachers at German vocational schools gave themselves an average rating of 3.4. However, they rated their ability to apply this knowledge in the classroom much lower, at 2.6. Knowing and doing had become disconnected. A trial across five European countries then tested whether training could close that gap. In the trial, 736 teachers in France, Ireland, Italy, Luxembourg and Slovenia were given a course and support. The results showed that their knowledge and ability to judge what a tool could do increased. However, their teaching did not change. They tried things and went back to what they had been doing before.

This is the knowledge–action gap. Courses produce knowledge, but changing practice needs accompaniment, trialing and colleagues to think with. **Transfer has to be designed, not hoped for.**

What teachers are actually doing

The most robust finding across countries is quietly surprising: t**eachers use AI intensively for themselves and hardly at all with their students**. In Germany, 52% use AI at least once a month. Only 29 % use it in lessons with a class in front of them. So the public debate about "AI in the classroom" is largely a debate about something that is not yet happening at scale.

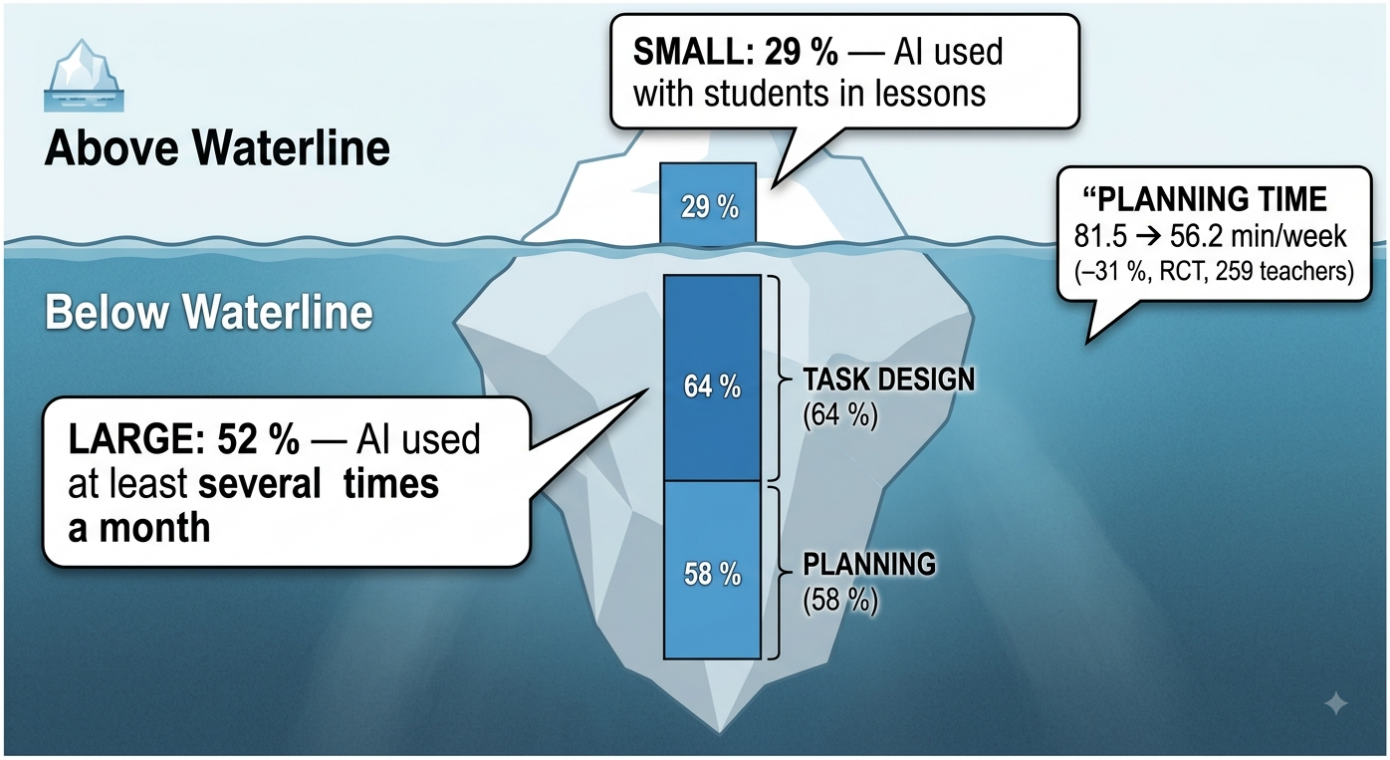

Fig. 2: Difference in AI use for preparation compared to usage in the classroom. Image created with GoogleGemini

What **is** happening is workload relief, and here the news is good. In a British trial, 259 teachers across 68 schools were given AI support for planning science lessons for 11- to 13-year-olds. Planning time fell **from 81 to 56 minutes a week**. When independent assessors reviewed the resulting materials, they found no drop in quality. Teachers' own estimates of the time they save across all subjects are even higher, at around six hours a week for regular users.

And one pattern is worth pausing on. In a large American survey, **teachers in schools with a clear written AI policy reported saving about a quarter more time than teachers in schools without one.** That is the teachers' own account rather than a measurement, and a survey cannot show which way the arrow runs. Well-organized schools may simply do both things well. But it is the opposite of what most people expect, and it is worth noting that nothing in the sources behind these points the other way: wherever this has been examined**, clear rules and heavy use go together.** -> 1.5 Compliant use of AI

Why the tool is not the point

If you wanted a ranking of the best AI tools, the research will disappoint you and then, on second reading, set you free. When researchers gather many separate studies and ask what separates the classrooms where AI helped a great deal from those where it barely helped at all, the answer is never the brand of software. It is what the students were asked to *do* with it.

There are three comparisons that really drive the point home.

- Students who used AI to transform something, e.g., rework a draft or build on their own attempt, gained about two and a half times as much as students who simply held a conversation with a chatbot.
- Students who used AI in pairs gained about 70% more than students who used it alone.
- AI was used to provide feedback on student work, and it's fair to say it made one of the biggest gains in this area. In contrast, using AI to incorporate gamification into lessons produced a gain too small to call proven..

Pedagogy beats technology, and by a wide margin. This is what happens every time somebody sorts the studies by what the students were actually doing. Two limits should be considered equally.

Good feedback is not the same as feedback that works. In one experiment, 70 university students wrote an argumentative essay and revised it a week later. They were assigned at random to three groups: one got feedback from an experienced teacher, one from a chatbot given a plain instruction, and one from the same chatbot prompted to reason through the essay step by step. Marked against the same rubric, the step-by-step chatbot produced the **best feedback of the three.** The feedback was better than the plain chatbot and better than the teacher. Then came the revisions, and **all three groups improved by much the same amount.** The best feedback in the study brought no more improvement than the lowest-rated feedback. **Better comments did not produce a better essay.**

That deserves a moment because it challenges a common assumption that better feedback leads to better learning. It does not, on its own. Feedback only does anything if it is **taken up**: read, understood, acted on, and understood well enough that the student could get it right unaided next time. **Uptake is the bottleneck, not quality**.

Two things appear to get in the way, and both are usable.

More isn't always better. In a second study, with 60 Year 11 students and 240 pieces of writing, the AI produced way more content-related comments than the teachers did. 333 against 241. It was the teachers' comments that the students acted on more often. When confronted with thirty precise corrections, a student might either stop diagnosing and start triaging or just comply. If a student follows the rules, they're basically saying that the revision was unnecessary. It's the satnav again, but on a different level. It gives you really detailed feedback, like 'Turn left in 200 meters' for your own writing.

Who says it matters. We're not really sure why teacher comments are taken up more often. This is all about interpretation, not measurement, but the most likely explanation is that feedback is seen as an act between people. What someone who knows what you did last term and will see what you do with it has to say is different from the same sentence generated on request.

What follows for practice. All the evidence we've seen on automated feedback shows that it works best when students take what they've learned in one task and apply it to a new task. It's least effective when they just copy and paste what they've learned straight from the text in front of them. And it's way more effective when you use it over a longer period of time, like a whole term,

rather than just in one go. **So, we shouldn't be asking whether the AI's feedback is good enough**. **We should be asking whether this student will ever have to use it again. If the answer is no, better comments are unlikely to help.**

And the same split shows up at scale. One team pooled 228 studies: AI clearly and consistently improved what students could produce, while any gain in their ability to plan, monitor and correct their own learning was small enough to have been chance. Programming shows it sharpest. Students with an AI assistant were measurably faster, and no better at the end than those without one. **Getting more done and getting better are not the same thing, and AI separates them.**

Europe as mirror

Across Europe the share of teachers using AI runs from **89 % in Czechia to 14 % in France.** Please consider that the surveys differ in method and timing, so read the spread, not the individual figures.

France is the instructive case. It built one of the continent's most ambitious state AI programs, an AI tutoring scheme in maths and French. Nevertheless, France still sits at the bottom of that table, as only 9 % of French teachers have received any training in AI, compared with roughly 38 % across comparable countries. **Providing tools is the easy half but qualification is the half that decides.**

Two countries have written the satnav problem into policy the way a driving school would: map first, device later. Norway recommended in June 2026 that generative AI remain essentially out of primary school and be taught only in lower secondary school, where teachers have been trained, and AI should be taught as a deliberate skill. Luxembourg stages it similarly. Note carefully: Norway issued a **recommendation**, not a binding ban, however the international press reported it.

The blind spot

One risk of AI is rarely named, and the image reaches it too: knowledge about the London city map **is not permanent.** A driver who follows the arrows for five years does not keep the city in her head. It fades, and she stops noticing the dug-up street. Your expertise is maintained by doing the very things AI is best at absorbing: writing the task, anticipating the misconception and drafting the explanation. **Where AI takes over the work through which judgment forms, relief turns into deskilling. In a German expert panel, 65 % of education experts identified this risk for the teaching profession, rising from 44 % the year before.**

Use the satnav. Just keep the map.

Revision #21
Created 2026-06-11 14:46:58 UTC by Jannik
Updated 2026-08-12 14:02:51 UTC by Weitzel