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.
- ***We argue that teaching about AI is the precondition for teaching with AI.*** You need to know the machine to check the answer it 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.*** In Germany, half of teachers use AI several times a month or more; of those, fewer than a third let pupils work with it in class.
- ***Pedagogy beats technology.*** Using AI to rework and build on one's own work beats using it only to practise or look things up; working in groups beats working alone; feedback beats gamification.
- ***Better feedback is not better learning.*** In one experiment, AI comments judged better than a teacher's led to revisions that were no better. What decides is whether feedback is *taken up.*
- ***AI raises output, not self-directed learning.***
- ***Clear rules and heavy use appear together.*** In a large US survey, schools with an AI policy had more teachers using AI – and so saved more time overall, not less.
- ***Guard your own map.*** *Relief becomes deskilling where AI absorbs the work that builds and maintains judgement.*

> *If you 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.

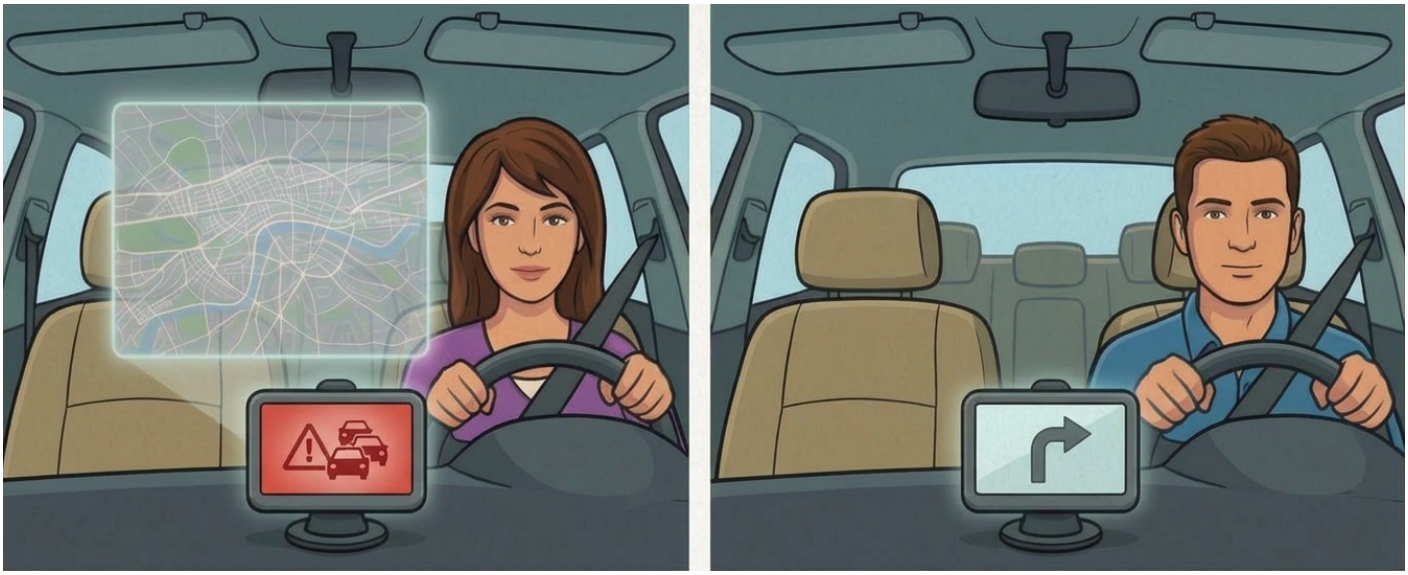


Fig. 1: "The outcome depends on the driver" (AI-generated image, Google Gemini)

On the left side the driving begins with a map in the head. The satnav adds what the driver cannot know. This is called "beneficial offloading". On the right the driver has no map in the head. The device is the only orientation for the driver. This is called "detrimental offloading".

Same satnav, two drivers, opposite outcomes. **The difference was never in the device.** We talk about more than an image. Studies of drivers find that people who rely heavily on satnav form a poorer mental map of their surroundings. A study over three years showed that heavier use went with a steeper decline in spatial memory. And psychologists have known for two decades that the same support which helps a beginner can get in the way of someone who already knows the subject. The same happens the other way around. 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 organised 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 practised with an unrestricted chatbot, another without. While they had it, the chatbot group scored **48 % higher on the 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 metres* 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.

Voices of the field

"Right now the tools students reach for are built to solve problems, not to teach, and this is the core difficulty in my work currently."

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.

In a study with 90 teachers at Bavarian vocational schools who had signed up for AI training, teachers rated their knowledge of AI at 3.4 out of 5 on average, but their ability to use it in their teaching much lower, at 2.6. Knowing and doing had come apart. A trial across five European countries then tested whether training could close that gap. 736 teachers in France, Ireland, Italy, Luxembourg and Slovenia took part. The teachers were split up in two groups. One group received an online course and support during the trial, the others received the same online course after completing the trial. The results showed that knowledge and the ability to judge what a tool could do, increased. But in most countries the regular classroom use of AI did not change: they tried tools out 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.**

Voices of the field

"The only drawback of this webinar is that we cannot try it out directly."

What teachers are actually doing

The most robust finding across countries is quietly surprising: **teachers use AI intensively for themselves and hardly at all with their students**. In Germany, 52 % of teachers use AI tools several times a month or more. But of those who use AI, only 29 % let their pupils work with it in class. So the public debate about "AI in the classroom" is largely a debate about something that is not yet happening at scale.

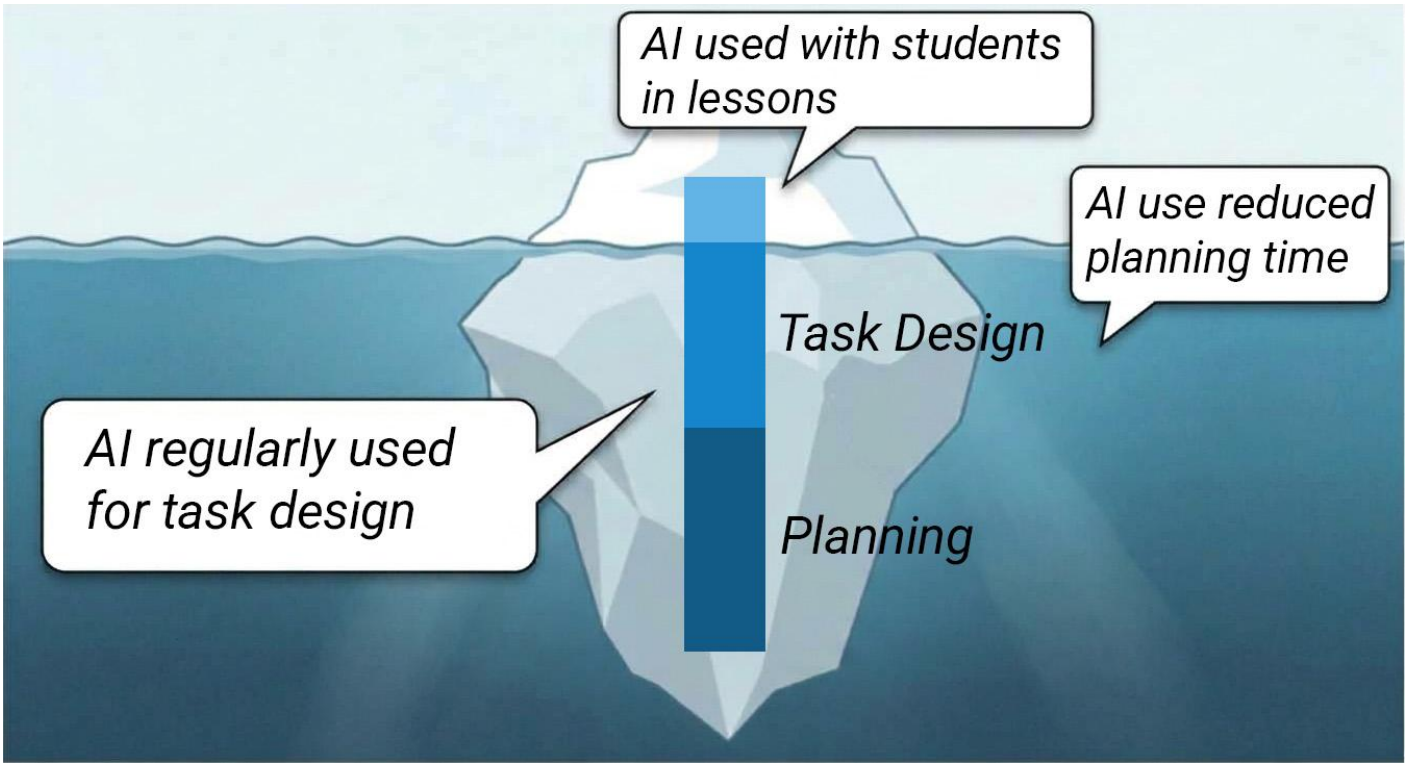


Fig. 2: "Difference in AI use for preparation compared to usage in the classroom" (AI-generated image, Google Gemini)

What **is** happening is workload relief, and here the news is good. In a trial in England, 259 teachers across 68 schools were given ChatGPT and a guide for planning science lessons for 11- to 13-year-olds. They spent **about 56 minutes a week on planning, against 81** in the comparison group. Independent assessors who did not know which group a lesson came from found no evidence of a difference in quality. In a large US survey, teachers who use AI at least weekly estimated saving almost six hours a week – their own estimate, not a measurement.

And one pattern is worth pausing on. In the same US survey, **schools with an AI policy had more teachers using AI – seven in ten against six in ten – and so saved about a quarter more teacher time in total.** For the individual teacher who uses AI every week, the difference

was small. That is the teachers' own account rather than a measurement, and a survey cannot show which way the arrow runs: well-organised schools may simply do both things well. But it is the opposite of what most people expect – **clear rules and heavy use appear together, not in opposition.** **1.4 Compliant Use of AI**

Voices of the field
"It's important to have it as a supplement, but definitely not let ChatGPT run the whole lesson."

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. The following bullet points summarise important research results. But all results build on only a few studies.

- In mathematics, students who used AI to rework a solution or to build on their own attempt gained about two and a half times as much as students who used it only to enhance ordinary tasks such as practising or checking answers.
- Students who used AI together in groups gained about 70 % more than students who used it alone.
- Where AI was used to give feedback on student work, gains were among the largest. Using AI to gamify lessons produced a gain too small to call proven.

Pedagogy is more important than the technology used. Where researchers have sorted the studies by what students were actually doing, the activity mattered more than the tool. Interestingly, the same pattern was found for the previous generation of digital tools in secondary schools. The results largely depend on whether teachers had been trained to use them.

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.**

This result challenges a common assumption that better feedback leads to better learning. Obviously 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.

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. But the students acted more often on the teachers' comments. 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 metres" for your own writing.

It matters who says it. 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. A pooled analysis of studies on automated writing feedback that was done before today's chatbots existed came to similar results. Feedback did more when students had to apply it to a new piece of writing than when they only revised the text in front of them, and more when it was used over several sessions rather than once. And one more thing should be said fairly: compared with no feedback at all, AI feedback does help. In a study by a German team with 459 upper-secondary pupils, AI-generated feedback modestly improved their revisions and their motivation.

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, surveys of teachers' AI use run from **about nine in ten in a Czech sample to about one in seven in France.** The surveys differ in method and timing: the Czech figure comes from a small 2025 survey of 317 teachers and counts any experience with AI whereas the French figure comes from the OECD's international teacher survey TALIS 2024, in which just under half of Czech teachers reported using AI in their teaching.

France built one of the continent's largest state AI programmes: adaptive practice software in maths and French, free for every pupil in the first year of upper secondary school. Nevertheless, French teachers use AI rarely and only 9 % of French teachers have received any training in AI, compared with 38 % on average across OECD countries.

Norway has written the satnav problem into its policy (at least as a recommendation) the way a taxi driving school would do. Learn with a map first, and afterwards add the device. In June 2026 Norway recommended that pupils in grades 1 to 7 should, as a rule, not be given access to AI. In

grades 8 to 10 AI may be introduced gradually and carefully, once teachers have the competence to guide it. In upper secondary school pupils should learn to use AI purposefully.

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 judgement forms, relief turns into deskilling.** In a German panel of e-learning experts, 65 out of 100 expected generative AI to cost people important skills, up from 44 the year before. In medicine the effect has already been measured: after AI support became routine in colonoscopy, doctors working without it detected fewer precancerous growths than before.

Use the satnav. Just keep the map.

Revision #39
Created 2026-06-11 14:46:58 UTC by Jannik
Updated 2026-10-20 17:02:21 UTC by Alexander