Didactics before tools

Redesigning teaching and assessment in the age of generative AI

Authors:

Eva Mårell-Olsson, Department of Education, Umeå University, eva.marell-olsson@umu.se

Tatjana Titareva, AI Policy Lab, Umeå University, tatjana.titareva@umu.se

Abstract

Research and policy have become good at naming the cognitive risks that generative AI (GenAI) poses to learning. Namely, detrimental cognitive offloading, metacognitive laziness, and the illusion of competence. What this literature does much less well is tell teachers what to *do* with that knowledge in their everyday practice. This article argues that alongside AI literacy, regulation and safeguards protecting educational stakeholders, one of the central challenges facing education in the age of GenAI is related to *didactic design.* When a finished text or answer no longer reveals what a student has understood, struggled with, revised, verified or learned, **teaching and assessment have to be redesigned so that the process of thinking is made visible and necessary again**. We argue that this redesign should be approached through the classic didactic questions: *why, what, how, when* and *for whom.* We connect these questions with Question Zero (Q0), an approach to responsible technology adoption, through the dimensions of *why, who, what, how and where.* Both didactic and Q0 perspectives and underlying questions move the starting point from the technology to educational purpose, stakeholders and the conditions under which the use of selected technology is well-designed and justified. We argue that current AI strategies, guidelines and institutional frameworks in education leave a conspicuous gap precisely here. We translate these didactic and Q0 perspective questions into a concrete, phased design for teaching and assessment, whose guiding principle is not to forbid cognitive offloading but to regulate it didactically. We also argue that making the learning process visible is not in itself a neutral act, and that its design must attend to who carries its costs. This challenge recurs across education systems globally and cannot be solved by top-down regulation alone. Instead, translating high-profile principles of ethical AI in education into pedagogically sound use of GenAI requires the contextual and professional judgement of educators.

Keywords: generative AI, didactics, Question Zero (Q0), cognitive offloading, assessment design, pedagogical design, professional judgement, AI in education

From cognitive risk to didactic response

Recent reports synthesise the evidence on AI and learning. A particularly useful contribution is the distinction between *beneficial* and *detrimental* cognitive offloading (Dunn & Risko, 2016; Risko & Gilbert, 2016; Lodge & Loble, 2026). AI can manage extraneous load and free working memory for the intrinsic work of learning, or it can bypass that intrinsic work altogether. Equally useful is the distinction between a *task-completion goal* and a *learn-the-concept goal* (Kalyuga & Plass, 2025). These distinctions give us a more precise language for what is at stake.

The problem is not simply that students might cheat, or that they might produce better answers than they could have produced alone. The more serious issue is that generative AI (GenAI) can interfere with the cognitive and metacognitive processes through which students learn (Fan et al., 2025). Students may complete the task without developing the knowledge, judgement or self-regulated learning capacity the task was meant to support. The most recent international synthesis reaches the same conclusion. The OECD (2026) finds that when students lean too heavily on GenAI, metacognitive engagement falls and task performance becomes misaligned with genuine learning.

The current empirical data and preliminary findings describe the risk well and give much less guidance on what teachers are supposed to do with this knowledge. How should they design teaching and assessment when the finished product no longer tells us very much about what the student has understood, revised, verified or learned? This is the question we want to place at the centre.

The pedagogical problem beyond AI literacy

Much of current education policy responds to GenAI by emphasising AI literacy and competence. We agree with their importance. But they do not constitute a holistic pedagogical response to cognitive outsourcing. Treating them as such risks diverting attention and resources from the difficult didactic work required (Weidlich et al., 2025). Knowing what a large language model is, and how to prompt it, does not by itself help a teacher decide how to assess understanding when fluent, confident output is available on demand. A recent article makes a parallel point: conventional, skills-based AI literacy frameworks remain anchored in functional competence and say little about the metacognitive and epistemic demands of probabilistic, opaque systems (Shapiro et al., 2026). That argument moves toward expanding literacy itself, into a critical, metacognitive and civic practice, whereas we argue the missing piece is the appropriate redesign of teaching and assessment.

In this opinion piece, we argue that the challenge of AI in education is fundamentally pedagogical rather than solely technological. The same tool produces opposite effects depending on the instructional design and the learner's goals. A systematic review of 67 empirical studies confirms the pattern. Generative AI's cognitive effects are not intrinsic to the tool but dependent on pedagogical framing, with scaffolded, inquiry-oriented use supporting higher-order thinking and unstructured use triggering cognitive offloading (Li et al., 2026). A risk taxonomy, however accurate, does not contain its own remedy. Naming "metacognitive laziness" tells a teacher what to fear, not how to handle it in the classroom.

The tension is structural rather than rhetorical. Cognitive offloading is, at its core, a way of reducing mental effort (Risko & Gilbert, 2016), and generative AI lowers the cost of such offloading further than any earlier technologies. Education rests on people working their way through uncertainty and complexity. This is not a superficial tension but a conflict between what the tool is built to optimise and what learning requires. Decades of research on *desirable difficulties* suggest that effort, struggle, and even the unpleasantness of thinking are often what produce durable learning (Bjork & Bjork, 2015; David et al., 2024), which is precisely the work that a tool designed to minimise effort invites students to skip. Early empirical work points the same way. In a study of 666 participants, more frequent use of AI tools was associated with weaker critical thinking, an association largely mediated by cognitive offloading (Gerlich, 2025).

If AI offers the fastest route to a finished product, then teaching has to make the *process* of thinking (i.e., explaining, revising, verifying, judging) more explicit. This is not a small adjustment. It touches the design of assignments, the role of feedback, assessment practices, the relationship between teacher and student, and students' motivation to engage in the difficult and sometimes frustrating work that learning requires.

Several strands of assessment scholarship that predate generative AI are especially relevant here. Work on assessment reform for an age of GenAI argues for distinguishing between assessment that secures and validates individual achievement and assessment that develops capability through authentic, process-rich tasks (Lodge et al., 2023; Dawson, 2021). Work on formative assessment and feedback shows how assessment can be designed to drive self-regulated learning rather than merely certify a product (Nicol & Macfarlane-Dick, 2006; Carless & Boud, 2018; Hathcoat et al., 2026). And the notion of *evaluative judgement* (i.e., the capacity to judge the quality of work, one's own and others') names precisely the capability that detrimental offloading erodes and that teaching now has to cultivate deliberately (Tai et al., 2018; Bearman et al., 2024). The OECD's synthesis points in the same direction. Effective integration, it argues, may require teachers to emphasise process, how students think and learn, rather than student output, and finds that hybrid systems pairing GenAI with explicit pedagogical models, such as structured tutoring or evidence-centred assessment design, show more promise than general-purpose chatbots (OECD, 2026). A 2025 UNESCO anthology reaches a similar conclusion, devoting a major section to assessment and arguing that standardised, output-focused testing is being destabilised by generative AI in favour of more human-centred, process-oriented forms of knowing (UNESCO, 2025b).

Making learning visible, but for whom?

There is a cautionary precedent worth recalling. Earlier efforts to make learning visible through digital tools did not automatically translate into changed teaching. A study of the introduction of mandatory personal development plans and digital documentation in Swedish compulsory schools found that, although the reform was meant to centre each student's learning, students' own goals had little effect on actual teaching. Responsibility for reaching them was largely devolved to the individual student, and the process drifted towards documentation, behaviour management and the production of the "model student" (Mårell-Olsson, 2012). Teachers, working under pressure, rationalised the task. For instance, teachers copied and adapted written assessments across students. The lesson is relevant to the current stage of GenAI adoption in education. Mandating a digital tool without redesigning teaching and assessment tends to create an administrative routine rather than reshape the learning process it was meant to support.

It is also important to be careful with the word 'visible'. Making the learning process visible is not a neutral or automatically benign act. The study showed how documentation practices can turn into surveillance, control and the disciplining of students towards a prevailing norm, a dynamic that Bernstein's distinction between visible and invisible pedagogy helps to name (Mårell-Olsson, 2012). Later studies of one-to-one computing classrooms found a similar dynamic, showing how teachers' enacted didactical design decisions carry symbolic power and control over students (Bergström et al., 2019). Thus, designing for process-visible learning has to attend to the 'for whom'. Visibility can support metacognition and self-regulation, or it can become a mechanism of compliance that weighs most heavily on the students with the least resources to navigate it.

The aim, therefore, is not visibility for its own sake but tasks designed so that the thinking serves learning rather than control, and so that students have a reason to understand the concept and apply it in the real world, rather than complete the task (see Lodge et al., 2023).

A didactic reframing: why, what, how, when, for whom

A further observation sharpens the design task. The evidence suggests that GenAI works best as a scaffold, not as a substitute. When it is used for Socratic questioning, feedback, alternative examples, concept explanations and support for self-regulation, it can support learning. When it is used to produce finished texts, solutions or analyses before the student has done their own cognitive work, the risk of shallow learning rises. A synthesis of research on curiosity and metacognition draws this distinction directly, warning that large language models' default interaction modes can act as a cognitive shortcut but when redesigned, they can instead become a partner for sustained epistemic development (Desvaux et al., 2026). The design question, therefore, is not AI or no AI, but when in the learning process AI is used, for what purpose, and with what degree of student responsibility. That reframing is what the didactic questions are built to handle.

We propose that AI in education be approached through the questions educators have long asked of any teaching decision. *Why* something should be taught, *what* learners need to know, *how* it should be taught, *when* in the course of learning, and *for whom*. These are the didactic questions of the Didaktik tradition (Klafki, 2000; Hopmann, 2007; Uljens, 1997). This tradition has already been extended to digital tools through, for example, the concept of digital didactical design, which provides an empirical framework for analysing how teachers' design decisions shape technology-mediated learning (Jahnke et al., 2017).

Approached this way, GenAI becomes one more factor in teaching design, to be shaped by the same questions rather than settled by a blanket rule:

- The *why* recovers the purpose of an assignment, which is often what makes offloading either harmless or harmful. This mirrors, at the level of the classroom, a broader argument that Europe's strategy should move from an AI first to a purpose first posture, placing societal benefit and human-centric values ahead of acceleration for its own sake (Dignum et al., 2025). What purpose first is to AI policy, the didactic why is to teaching.
- The *what* forces a distinction between knowledge that must be built in the student and knowledge that may legitimately be offloaded.
- The *how* and *when* concern where in a learning sequence friction is desirable and where it is merely unnecessary.
- The *for whom* is not incidental. The cognitive risks of AI fall hardest on novices and on students with weaker self-regulation, raising the prospect of a widening equity gap

(Bastani et al., 2025). It is the same population, and the same equity concern, raised by the caution about visibility above. Making the process visible can support these students or, if designed carelessly, weigh most heavily on them. This is also why frameworks aimed only at upgrading teachers' technical competence, important as they are (Redecker, 2017; UNESCO, 2024a, 2024b), need to be complemented by frameworks for pedagogical and content-specific judgement (Mishra & Koehler, 2006; Mishra et al., 2023).

Next, we introduce the Question Zero tool from the policy world with five overarching questions-dimensions related to thoughtful and responsible AI adoption and in what ways they overlap and complement the didactic questions.

Question Zero: why, who, what, how, where

Question Zero (Q0) from the policy world asks: *"Under what conditions should an AI system be adopted, if at all?"* (Dignum et al., 2026; Titareva et al., 2026). Its purpose is to resist an "AI First" logic in which the technology is perceived as the *only* next step. And problems are identified later for the technology to solve. Instead, the problem and available alternatives (why?), affected stakeholders (who?), type of technology to solve a specific problem at hand (what?), and conditions of adoption (how and where?) should be considered before an AI tool is selected as the solution.

In education, a similar mistake is to begin with the assumption that generative AI will be used and then impose the decision on teachers and students to use it without pedagogically sound guidance. A didactic approach requires an earlier question: why we are using the tool and what educational problem are we trying to solve?

This is where Q0 and the classical didactic questions, coming from different disciplines and angles, become connected. Q0 is a governance question about responsible and thoughtful pre-adoption stage; the didactic *why* is a question about educational purpose. Both questions move the focus from technology as their starting point to the purpose. Q0 requires organisations to describe honestly the existing problem (not the solution/tool), consider human and non-AI alternatives, and justify why AI is appropriate before adopting it. The didactic *why* for teaching asks what an activity is intended to achieve (desirable teachers and learners' goals) before the stakeholders decide what tools to use.

The connection between didactic questions and Q0 dimensions moves beyond *why*. Q0 operationalises responsible adoption through five iterative dimensions: w*hy, who, what, how* and w*here*. These dimensions include: motivation for using the tool starting from the real problem identification and alternatives, affected stakeholders, the type and implications of the AI system, the adoption process, and infrastructure and monitoring. Didactic questions ask what happens when technology is imposed on teaching practices: what knowledge has to remain with the learner, when assistance should be provided, how much cognitive work may be offloaded, and for whom a particular design encourages or limits learning. The two approaches and questions are complementary and help make informed and weighted decisions prior to technology adoption in core activities, including in teaching and assessment.

Table 1

Didactic Questions vs. Q0 Dimensions

Didactic Questions	Question Zero dimensions
Why?	Why?
What?	Who?
How?	What?
When?	How?
For whom?	Where?

Why top-down regulation is not enough

It is tempting to respond to the resulting uncertainty with detailed, centralised rules, which is important for such complex systems and environments as education. Indeed, the lack of concrete guidance does create a large space for interpretation, which will likely produce considerable variation in how AI is understood and used across classrooms and institutions (Mårell-Olsson, in press). However, education cannot be governed as if teaching was mainly a compliance task. Teachers need pedagogical freedom and the exercise of professional evaluated judgement, because the right use of AI is contextual. It depends on the subject, the stage, the goal and the student.

The current high-profile policy documents vividly demonstrate it. High-level, rights-based guiding documents such as UNICEF's "Guidance on AI and Children" (2025) and UNESCO's "AI and education: protecting the rights of learners" (2025a) are valuable precisely because they are general enough to apply across very different contexts. However, their authors are explicit that these guidelines must remain high-level, and they do not and cannot be easily applied at the level of each classroom. Even the European Commission's dedicated, recently updated "Ethical guidelines on the use of AI and data in teaching and learning for educators" (European Commission, 2026), which now incorporates the AI Act and GDPR and offers core principles, guiding questions and classroom scenarios, is framed around ethical and responsible use and risk awareness rather than around the teaching redesign that the cognitive-offloading problem demands. The same is true of most national AI strategies and sectoral guidelines. They identify principles and risks. They do not, and arguably should not, prescribe teaching design. That work has to be done with and by teachers, not to them.

From high-profile principles to practice: designing teaching and assessment in GenAI times

Assessment has to begin by measuring process, judgement and independence, not only the finished product. The research and policy debate on AI and assessment points towards *combinations* of AI-free and AI-integrated components (Lodge et al., 2023; Dawson, 2021). AI-free components are needed to secure basic individual understanding. AI-integrated components are needed to assess how a student uses AI critically, ethically and in a subject-responsible way.

These are complementary, not alternatives. Below we propose a framework for putting didactic/Q0 questions into practice in pedagogical settings in the age of generative AI.

We propose designing teaching in phases:

1. **Stage 1. Own understanding**, without AI (i.e., securing initial cognitive work): The purpose is to protect the cognitive work through which initial understanding is built. The student must first read, think, formulate and test ideas on their own.
2. **Stage 2. AI as dialogue partner**, critic or feedback resource (i.e., extending and challenging thinking): GenAI is used only to extend thinking rather than replace it. AI is introduced to question, challenge, exemplify and provide feedback.
3. **Stage 3. Verification** (i.e., developing disciplinary and epistemic judgement): AI use is conditional. The student must compare GenAI tool's output against course literature, empirical evidence, theory or their own data, thereby turning passive acceptance into active judgement.
4. **Stage 4. Accountability** (i.e., making the use of AI transparent and reflective): The student reports what AI was used for, what was accepted, what was rejected, and why.

Stage 5. **Monitoring:** Assessment can then ask a different set of questions: Can the student explain this *without* AI? Can the student use AI *without* losing subject-matter judgement? Can the student identify errors, uncertainties and limitations? Can the student defend their choices, orally or in writing? Can the student show how their understanding developed over time?

Table 2

5-Stage AI Pedagogy Framework (Mårell-Olsson & Titareva, 2026)

Stage	Focus & Purpose	AI Status	Key Student & Instructor Actions
1. Own Understanding	Secure initial cognitive work	**No AI**	Student reads, thinks, formulates, and tests initial concepts independently.
2. Dialogue & Challenge	Extend thinking	**Active AI Partner**	AI acts as critic / Socratic dialogue partner to question and exemplify.
3. Verification	Develop epistemic judgement	**Conditional AI**	Student cross-references AI outputs against peer-reviewed literature, theory, and data.

4. Accountability	Transparency & reflection	**Transparent AI**	Student documents AI prompts, accepted/rejected output, and rationale.
5. Monitoring	Assessment of core mastery	**Evaluated Mastery**	Assessment tests independent defence, error detection, and student growth over time.

We are not claiming to offer a novel solution to designing teaching and assessment for GenAI reality. A recent framework for higher education proposes similar principles, preserving cognitive friction, positioning AI as a provisional thinking partner, embedding evaluation throughout, and sequencing AI-free with AI-mediated phases (Vendrell & Johnston, 2026). Wazan (2026) builds an assessment design and examination system around the same logic. Students must commit to an initial answer *before* consulting GenAI, then interrogate the model's output (including deliberately flawed responses) and justify a refined answer, scored by a rubric that assesses reasoning as a *process* rather than a product. That two such proposals, developed from cognitive psychology, critical pedagogy and epistemology tradition, converge on the same sequence is significant in itself. What we add is the didactic account of *why* it works, the why, what, how, when and for whom, its extension across schooling as well as higher education, and the caution developed above that making the process visible is not automatically benign.

To illustrate our framework in practice, consider a written assignment that requires students to analyse a problem using course material. In the first phase, students draft their analysis in class, without AI, so that the initial reading and reasoning are their own. In the second stage, students use AI to interrogate their draft, asking it to challenge the argument, surface counterexamples or offer alternative framings, and then deciding what to keep. In the third phase, students check every AI-generated claim, source and example against the literature and the evidence, and correct what does not hold. In the fourth stage, students submit a short account of how AI was used, what they accepted or rejected and why, and are examined on whether they can defend the analysis without it. In the fifth stage, the educator assesses how students can explain their analysis without AI, show critical thinking and demonstrate the growth of their own thought and knowledge over the process of completing the assignment.

These phases are a design logic rather than a fixed timetable. They may recur within a longer task, and where friction is most valuable depends on the students and the subject. A novice may need a longer, more protected first phase, while an advanced student may begin closer to verification.

We believe that, rather than forbidding cognitive offloading, education in the future must *regulate it pedagogically*. AI may reduce extraneous load. It must not be allowed to take over the load that builds understanding. This principle, drawn from the distinction between beneficial and detrimental offloading, is helpful in guiding teaching and assessment design in GenAI times.

Implications for professional development and a different kind of guidance

We offer two crucial implications. First, education systems face a very large need for professional development, from schools to higher education globally. This goes beyond AI literacy. Teachers need practical pedagogical support in designing teaching and assessment that reduces detrimental offloading and metacognitive laziness, while still allowing beneficial offloading, and in helping students understand why the goal is to learn the concept and its applicability to the real world, not merely to complete the task (see Lodge et al., 2023). Policy recommendations are starting to reflect this need. The Council of the European Union's 2026 "Conclusions on teachers in the era of artificial intelligence (AI)" place teachers at the centre, recognising that AI may reshape how they design, deliver and assess instruction and calling for guidance and supportive frameworks for their continuous development (Council of the European Union, 2026). Characteristically, though, such instruments invite and recognise rather than specify what that development should contain, which is the gap we attempt to cover in this article.

Second, the kind of guidance that policy produces needs to pivot. The most useful instruments for teachers and classrooms will not be longer lists of risks or prohibitions, but resources that help teachers translate risks into concrete decisions about assignment design, feedback, and assessment (working examples, design patterns, and exemplars of process-visible tasks) while preserving professional judgement. Better AI tools may help at the margins, such as Socratic tutoring models, tools designed to scaffold metacognition rather than provide complete answers, and teacher-facing applications that augment rather than replace an expert (Wang et al., 2024; OECD, 2026). The OECD frames this as a choice between replacement, complementarity and augmentation, and argues that augmentation (e.g., teacher and AI refining each other's work, while professional judgement is preserved) holds the most promise (OECD, 2026). But tools are not the fundamental issue. The fundamental issue is how generative AI changes the design of teaching and the assessment of students' knowledge.

Conclusion

Many current AI policies and reports have grown adept at identifying the risks generative AI poses to learning. They are useful but give teachers very little practical support in translating those risks into teaching and assessment design. We argue that closing that gap is the central educational task of this moment, and it is a shared task across education systems globally. Just as AI policy is being urged to put purpose before acceleration (Dignum et al., 2025), **education should put didactics before tools**. The argument is not merely that didactics matter, but that the didactic questions can be turned into a concrete, phased design for teaching and assessment, one whose guiding principle is to regulate cognitive offloading rather than forbid it. Closing the gap, therefore, depends a lot on the design work, and on investing, at scale, in the professional judgement of educators.

In conclusion, earlier we introduced Question Zero (Q0) as a way of asking the right questions before adoption: "Under what conditions should an AI system be adopted, if at all?" Q0 includes dimensions of why, who, what, how and where. What real problem are we trying to solve? Who might be affected? What kind of technology, if any, is appropriate to solve a problem at hand? How should it be adopted and governed? And where, and under what conditions should it operate? The didactic questions ask: why, what, how, when and for whom? The two approaches converge on one overarching premise: the adoption of the technology does not start with the tool, instead it starts with the purpose, i.e., answering why and what do we want to achieve? Both invite **us** to think from a perspective of purpose, context, people and the goals and conditions in which the tool is planned to be applied.

References

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, *122*(26), Article e2422633122. https://doi.org/10.1073/pnas.2422633122

Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgement for a time of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, *49*(6), 893–905. https://doi.org/10.1080/02602938.2024.2335321

Bergström, P., Mårell-Olsson, E., & Jahnke, I. (2019). Variations of symbolic power and control in the one-to-one computing classroom: Swedish teachers' enacted didactical design decisions. *Scandinavian Journal of Educational Research*, *63*(1), 38–52. https://doi.org/10.1080/00313831.2017.1324902

Bjork, E. L., & Bjork, R. A. (2015). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (2nd ed., pp. 59–68). Worth.

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Council of the European Union. (2026). *Council conclusions on teachers in the era of artificial intelligence (AI)* (9003/26). https://data.consilium.europa.eu/doc/document/ST-9003-2026-INIT/en/pdf

David, L., Vassena, E., & Bijleveld, E. (2024). The unpleasantness of thinking: A meta-analytic review of the association between mental effort and negative affect. *Psychological Bulletin*, *150*(9), 1070–1093. https://doi.org/10.1037/bul0000443

Dawson, P. (2021). *Defending assessment security in a digital world: Preventing e-cheating and supporting academic integrity in higher education*. Routledge. https://doi.org/10.4324/9780429324178

Desvaux, C., Abdelghani, R., Oudeyer, P.-Y., & Sauzéon, H. (2026). *Curiosity and metacognition: Towards a unified framework for learning and education in the age of AI* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2604.25648

Dignum, V., Carli, R., Ericson, P., Titareva, T., & Tucker, J. (2025). *"AI first" to "purpose first": Rethinking Europe's AI strategy*. AI Policy Lab (AIPEX), Umeå University. https://doi.org/10.63439/LPOU6506

Dignum, V., Carli, R., Dahlgren Lindström, A., Ericson, P., Titareva, T., & Tucker, J. (2026 June). Question Zero for Explainability and Vice Versa: The Case of the EU's AI First Strategy. In *International Conference on Human-Computer Interaction* (pp. 17-31). Cham: Springer Nature Switzerland.

Dunn, T. L., & Risko, E. F. (2016). Toward a metacognitive account of cognitive offloading. *Cognitive Science*, *40*(5), 1080–1127. https://doi.org/10.1111/cogs.12273

European Commission. (2026, May 12). *Guidelines on the ethical use of artificial intelligence and data in teaching and learning*. European School Education Platform. https://school-education.ec.europa.eu/en/discover/publications/guidelines-ethical-use-ai-and-data-teaching-and-learning

Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, *56*(2), 489–530. https://doi.org/10.1111/bjet.13544

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, *15*(1), Article 6. https://doi.org/10.3390/soc15010006

Hathcoat, J. D., Slotnick, R., & Miller, W. (2026). The End of Assessment? Disruption and Transformation in the Age of AI. *Research & Practice in Assessment*, *21(3), 5-26* (Special Issue–AI in Assessment).

Hopmann, S. (2007). Restrained teaching: The common core of Didaktik. *European Educational Research Journal*, *6*(2), 109–124. https://doi.org/10.2304/eerj.2007.6.2.109

Jahnke, I., Bergström, P., Mårell-Olsson, E., Häll, L., & Kumar, S. (2017). Digital didactical designs as research framework: iPad integration in Nordic schools. *Computers & Education*, *113*, 1–15. https://doi.org/10.1016/j.compedu.2017.05.006

Kalyuga, S., & Plass, J. L. (2025). *Rethinking cognitive load theory*. Oxford University Press. https://doi.org/10.1093/9780190078539.001.0001

Klafki, W. (2000). Didaktik analysis as the core of preparation of instruction. In I. Westbury, S. Hopmann, & K. Riquarts (Eds.), *Teaching as a reflective practice: The German Didaktik tradition* (pp. 139–159). Lawrence Erlbaum Associates.

Li, C., Cui, H., & Hagedorn, L. S. (2026). The cognitive impact of ChatGPT in higher education: A systematic review of critical and creative thinking outcomes. *Computers and Education: Artificial Intelligence*, *10*, Article 100571. https://doi.org/10.1016/j.caeai.2026.100571

Lodge, J. M., Howard, S., Bearman, M., Dawson, P., & Associates. (2023). *Assessment reform for the age of artificial intelligence*. Tertiary Education Quality and Standards Agency (TEQSA). https://www.teqsa.gov.au/sites/default/files/2023-09/assessment-reform-age-artificial-intelligence-discussion-paper.pdf

Lodge, J.M., Loble, L. (2026). Artificial intelligence, cognitive offloading and implications for education. University of Technology Sydney. https://doi.org/10.71741/4pyxmbnjaq.31302475

Mårell-Olsson, E. (2012). *Att göra lärandet synligt? Individuella utvecklingsplaner och digital dokumentation [Making learning visible? Personal development plans and digital documentation]* [Doctoral dissertation, Umeå University]. DiVA. http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-59420

Mårell-Olsson, E. (in press). AI literacy, governance, and equity in Swedish education policy. In T. Titareva, A. Csonska-Stambekova, & E. Mårell-Olsson (Eds.), *AI ethics in global education: Building a future grounded in local realities*. CRC Press.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Mishra, P., Warr, M., & Islam, R. (2023). TPACK in the age of ChatGPT and generative AI. *Journal of Digital Learning in Teacher Education*, *39*(4), 235–251. https://doi.org/10.1080/21532974.2023.2247480

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. https://doi.org/10.1080/03075070600572090

OECD. (2026). *OECD digital education outlook 2026: Exploring effective uses of generative AI in education*. OECD Publishing. https://doi.org/10.1787/062a7394-en

Redecker, C. (2017). *European framework for the digital competence of educators: DigCompEdu* (Y. Punie, Ed.). Publications Office of the European Union. https://doi.org/10.2760/159770

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Shapiro, H., Souto-Otero, M., & Watermeyer, R. (2026). Metacognitive AI literacy: Going beyond the AI skills gap agenda. *Learning, Media and Technology*. https://doi.org/10.1080/17439884.2026.2652638

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education, 76*(3), 467–481. https://doi.org/10.1007/s10734-017-0220-3

Titareva T., Tucker J., Carli R., Movchan V., Dignum V. (2026). *Question Zero: Why Responsible AI Begins Before AI Adoption.* AI Policy Exchange Forum (AIPEX). https://doi.org/10.63439/WCEZ7337

Uljens, M. (1997). *School didactics and learning: A school didactic model framing an analysis of pedagogical implications of learning theory*. Psychology Press. https://doi.org/10.4324/9780203304778

UNESCO. (2024a). *AI competency framework for teachers*. https://doi.org/10.54675/ZJTE2084

UNESCO. (2024b). *AI competency framework for students*. https://doi.org/10.54675/JKJB9835

UNESCO. (2025a). *AI and the future of education: Disruptions, dilemmas and directions*. https://doi.org/10.54675/KECK1261

UNESCO (2025). *AI and education: protecting the rights of learners.* https://doi.org/10.54675/ROQH4287

UNICEF. (2025b). *Policy guidance on AI for children* (Version 3.0). https://www.unicef.org/innocenti/reports/policy-guidance-ai-children

Vendrell, M., & Johnston, S.-K. (2026). Scaffolding critical thinking with generative AI: Design principles for integrating large language models in higher education. *Computers and Education: Artificial Intelligence*, *10*, Article 100572. https://doi.org/10.1016/j.caeai.2026.100572

Wang, R. E., Ribeiro, A. T., Robinson, C. D., Loeb, S., & Demszky, D. (2024). *Tutor CoPilot: A human-AI approach for scaling real-time expertise* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2410.03017

Wazan, A. S. (2026). *Strategies for creating uncertainty in the AI era to trigger students' critical thinking: Pedagogical design, assessment rubric, and exam system* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2602.00026

Weidlich, J., Gašević, D., Drachsler, H., & Kirschner, P. (2025). ChatGPT in education: An effect in search of a cause. *Journal of Computer Assisted Learning*, *41*(5), Article e70105. https://doi.org/10.1111/jcal.70105