



ACTION PAPER · MAY 2026

AI in K-9 schools

A short action paper for teachers and school leaders — a distillation of the research overview "Artificial Intelligence in Danish K-9 Schools"

TEACHERS & LEADERSHIP

DIGEST

MAY 2026

© 2026 Martin Hasbirk Brandt

Contents

About this document	3
1. The most important finding	4
2. What students are actually doing right now	6
3. What teachers get out of it	7
4. The biggest risk: cognitive offloading	8
5. Recommendations differentiated by age group	9
6. How to get started — tomorrow, this week, this month	11
7. How to test whether an AI tool is guided enough	12
8. When you suspect cheating	13
9. Parent communication	14
10. For school leadership and municipality	16
11. Caveats to know before citing this onwards	17
Appendix A — Reading guide to the research overview	18
Appendix B — Glossary	19

About this document

This is the short version. It is the document you bring to a staff meeting, send to your principal, or read before making a decision. It covers the essentials from the research overview *Artificial Intelligence in K–9 Schools — What Western research actually says about AI in grades 1–9, and what a Danish teacher can do with it* (May 2026). Cross-references to the full overview are noted throughout so you can look up a point when it needs to be substantiated.

This is not a neutral summary. It is a filtered selection, as the full overview is. Where I have had to choose between evidence and conviction, I have chosen evidence. Abbreviations and technical terms not explained in the text are collected in Appendix B at the end.

A note on terminology: When I write *guided AI*, *scaffolded AI* and *pedagogically framed AI*, I mean the same thing — an AI that is built to guide rather than answer, to ask questions rather than deliver complete solutions, and where the teacher retains oversight of the student’s conversation. The opposite is *free AI* or *AI without guidance* — typically ordinary ChatGPT access with no pedagogical constraints. The document is centred on Denmark, where the Folkeskole covers compulsory schooling for ages 6–15 (grades 0–9). I use “K–9” as a shorthand for the same age range throughout.

Who should read what?

- **School leader or municipal decision-maker** wanting a quick overview: Start with **section 10** (school leadership and municipality) and **section 11** (caveats). Then return to **section 1** (the most important finding).
- **Teacher in grades 1–4:** Read **section 5** and **section 6** (in that order). The other sections are context.
- **Teacher in grades 5–6:** Read **sections 5, 6, 7** (the tool test) and **9** (parent communication).
- **Teacher in grades 7–9:** Read **sections 5, 6, 7, 8** (cheating) and **9** (parent communication).
- **Everyone:** Read **section 11** (caveats) before citing this document elsewhere.

1. The most important finding

It is the design of the AI tool that determines the outcome — not its mere presence. The studies that succeed are those where the AI is built with a pedagogical framework: short replies, Socratic questions, one step at a time, no complete solutions. The studies that fail are those where students have free access to ChatGPT and let it do the work for them.

Three examples close to the Danish K–9 context:

- **Stanford Tutor CoPilot (Wang et al., 2024)** — pre-registered RCT with 900 tutors and roughly 1,800 American K-12 students. The AI's primary user here is *the adult tutor*, not the student. The tool nudged the tutor towards asking guiding questions. Students scored four percentage points higher on academic measures, nine percentage points higher when the tutor had originally received low ratings. Cost: about 20 dollars per tutor per year. The important point is the change in principle: AI as an amplifier of the master-apprentice relationship, not as a standalone student tutor.
- **Kestin et al. (Harvard, *Scientific Reports*, June 2025)** — RCT, n=194 university students, physics. A GPT-4-based tutor with tight pedagogical guidance outperformed traditional classroom teaching. The students learned more than twice as much in less time (median 49 versus 60 minutes). The probability that the result was due to chance is less than one in 100 million. Caveat: adult students, not children.

The one study to remember: Bastani

Bastani et al. (PNAS, 2025) — RCT, roughly 1,000 students in grades 9–11, Turkish mathematics. Three experimental arms: control, GPT-Base (free ChatGPT access), and GPT-Tutor (with pedagogical safeguards).

During practice, accuracy rose 48 per cent (Base) and 127 per cent (Tutor). On the subsequent exam *without AI*:

- **GPT-Base** scored 17 per cent *lower* than the control group.
- **GPT-Tutor** ended up roughly *on par with* the control group — no harm, but no lasting gain.

Students in the Base arm copied solutions without learning, and **they did not notice the loss themselves**. That is the single most important finding for Danish K–9 policy.

Take the three studies together, and you have the single most important point for Danish K–9 policy in one line:

AI as a teacher's tool works. AI as a student's crutch causes harm. AI as a guided collaborator can be used without harm — but do not expect it on its own to produce lasting learning for the student.

That is the nuance section 5 unfolds by age group. This is not a ban on a student ever touching the AI — it is a ban on the *free, unguided* use, which is what dominates at home and in students' spare time.

See the research overview, section 01 “Effects on learning outcomes”, for the full picture, including Eedi/LearnLM (UK, ages 13–15), the World Bank’s Nigeria experiment, and Kosmyna et al.’s EEG study from MIT Media Lab.

2. What students are actually doing right now

Danish K–9 data are scarce. The best available data come from our Nordic neighbours:

- **Norway (Spørringen 2025, the national school survey):** 65 per cent of leaders at the primary level (grades 1–7) and 89 per cent at the lower secondary level report that students use AI.
- **Sweden (Skolverket, December 2024, n=529 teachers):** roughly 40 per cent of teachers in the *grundskolan* use AI services. Only one in six allows students to do so.
- **EVA (the Danish Evaluation Institute, 2025, n=1,411):** two thirds of final-year upper-secondary (gymnasium) students have used AI in a way they themselves classify as cheating. That is *gymnasium*, not K–9 — but it is the most rigorous Nordic dataset on student AI behaviour, and it is what a Danish lower-secondary teacher must come to terms with.

The peer effect is strong. Norwegian students begin using AI earlier and more widely than their Swedish and Finnish counterparts. The students at your school are no exception.

See the research overview, section 05 “Policy frameworks in the Nordics and the West”, for full country figures.

3. What teachers get out of it

This is the area where the evidence most clearly points to clear gains. AI as a tool for teachers saves time without compromising quality, provided it is introduced with the tool and the training together. But it is not without concern — particularly from parents.

- **EEF × NFER (Education Endowment Foundation, UK, 2024):** teachers who used ChatGPT with an accompanying toolkit reduced planning time by 31 per cent (25.3 minutes per week). An independent panel could see no difference in teaching quality.
- **RAND (USA, 2024–25):** teacher use has doubled from 25 to 53 per cent in a year. Local teacher training lags 12–24 months behind. 61 per cent of parents and 22 per cent of local school leaders are concerned about AI's effect on critical thinking.
- **German–Danish teacher comparison (Tandfonline 2025, n=532):** German teachers have greater confidence in their own AI skills. The decentralised Danish system and the strong professional autonomy of Danish teachers shape how AI is being used — in Denmark, it happens in isolated pockets among individual teachers, not as coordinated, school-wide practice.

The important finding from RAND is that teachers use AI for *preparation*, not direct teaching. They keep their professional authority over the subject matter. AI takes over preparation, not the classroom.

See the research overview, section 03 “Effects on teachers”.

4. The biggest risk: cognitive offloading

The research does not say “AI harms cognition” as a blanket statement. It says something more precise: **AI shifts cognitive load, and where it lands depends on how the task is built.** Under free access, students’ depth of learning, retention and argument quality all decline. Under scaffolded use, where the AI is forced to ask questions rather than answer, gains appear.

The most worrying finding is that students do not notice the loss themselves. The Bastani study (PNAS 2025), the Kosmyna study (MIT 2025) and Bjorklund et al. (BJET 2024) point in the same direction: students believe they are learning, while their actual ability declines. That is the narrative we, as teachers, have to face — and it is what makes the practical choices in section 5 matter.

See the research overview, section 04 “Cognitive risks”.

5. Recommendations differentiated by age group

This is the core of what I would base my decisions on if I were sitting on a school leadership team or in a municipality in 2026 and had to act on the existing evidence. It is not a universal remedy.

An honest caveat before the recommendations: The evidence-based design principles I draw on (Kestin, Stanford CoPilot, Eedi) come from studies of adult students, adult tutors, and pupils aged 13–15. When I extend those principles to grades 5–6, it is a deliberate extrapolation. It is necessary because no rigorous RCTs exist for early and middle primary (see caveat 1). If you prefer caution, run grades 5–6 on the grades-1–4 model until K–9 evidence appears. That is a defensible professional choice.

Grades 1–4 (ages roughly 6–10)

Teacher-mediated AI only. In practice: students have no independent access to generative AI in school. The teacher may very well show students what the AI produces — for example, read aloud an AI-generated draft, display it on the smartboard as material for discussion, or mention that the differentiated tasks were prepared with AI's help. Students see the product, but they do not operate the machine.

Use AI for the *teacher's* work: lesson preparation, differentiation, reading scaffolding, drafts of individual action plans. Only on the condition that GDPR and the STIL guidance are observed.

The reasoning is threefold: UNESCO's age limit of 13, no real RCT evidence for AI tutoring in early primary, and the highest risk of cognitive offloading at a stage when foundational skills are still forming.

Threshold for reassessment: If large RCTs emerge for the youngest grades, you should revisit your school policy. The UK EEF's *Aila trial* (an AI lesson planner from Oak National Academy) running in 2024–26 will be the first to test this seriously. School leadership should make a note: the first results are expected in autumn 2026.

Grades 5–6

Supervised, scaffolded work with AI literacy. Use a Danish tool with pedagogical guidance built in. Two realistic options:

- **SkoleGPT** (skolegpt.dk) — open, free, GDPR-compatible. Developed by Future Classroom Lab in collaboration with CFU/Copenhagen University College and DBC Digital. Won the Danish Sprogprisen language prize in 2023. Strongest if you value openness and full control of the infrastructure.
- **Miira** (miira.dk, Pathfindr ApS in Roskilde) — commercial, a mature product, ISAE 3000 Type II audited (an international auditing standard for data processors). The school chooses the underlying model. Strongest if you value a polished pedagogical tutor with EU-only infrastructure and a separate teacher assistant.

The principled first choice is SkoleGPT, because it is open and free (see section 10). Miira is the mature commercial alternative if you need a tested user interface and are willing to pay for it.

Present AI as a *cognitive collaborator*, not as an answer machine. That must be built into the tool — you should not have to enforce it manually. Follow the design principles from Kestin/Stanford/Eedi: short answers, no complete solutions, Socratic questions, and the teacher being able to see the student’s conversation. Build on Generation AI (a six-year Finnish research project on AI literacy for younger students, run by the universities of Helsinki, Oulu, and Eastern Finland) and on the educational materials from Aarhus University’s Center for Computational Thinking and Design.

How do you know whether your tool is guided enough? See **section 7** for a short checklist.

Grades 7–9

Integrate AI in a way that preserves the student’s cognitive effort. Apply what I will call the *Bastani pattern* (from the study in section 1): students may use AI during exercises, but they are *always* assessed without it. This covers both summative assessments (graded marks, level placements) and documented ongoing assessments. The teacher may still give formative feedback that builds on a conversation in which the student used AI along the way.

Teachers should talk openly with students about the offloading mechanisms — why it is easy to let the AI do the work, and what the student misses out on as a result. The task design assumes AI is available, so the weight shifts towards oral defences, process portfolios (collected documentation of the student’s working process), and supervised writing as in the FP9 (Danish 9th-grade exit exams).

Skolverket’s recommendation to avoid unsupervised homework is sensible here. Teach AI literacy explicitly: algorithmic bias, hallucination, prompt design, and source criticism.

See the research overview, section 09 “Recommendations”, for the full version, including specific tool choices.

6. How to get started — tomorrow, this week, this month

This is the practical part. The short version: Start with yourself.

If you teach grades 1–4

Tomorrow: Students get no new access to ChatGPT or similar — the starting point is that they do not have it in lessons. Spend 15 minutes trying SkoleGPT yourself with a specific task from next week's plan: ask it to produce four versions of a reading task at different levels.

This week: Ask your year-group colleagues to show each other one concrete example of something they have done with AI for preparation. Twenty minutes is enough.

This month: Raise it at a team or year-group meeting: "Which tasks can the AI produce a good first draft of, so we save time without compromising the academic quality?"

If you teach grades 5–6

Tomorrow: Write down for yourself: In which two academic situations next week does it make sense to try SkoleGPT or Miira *together with the class* as a questioning tool? These should be situations where students can see your screen and where you discuss the output aloud.

This week: Check with the school leadership which tool is approved for use at your school. If none is approved, hold off on using it with students — and use the time to learn the tool yourself. Run the checklist in section 7 on the tool you are considering.

This month: Run one whole-class lesson where the class uses the approved tool as a thinking partner, not as an answer machine. Ask the students to describe what the AI got right and what it got wrong.

If you teach grades 7–9

Tomorrow: Have a conversation with your students for 5–10 minutes: "What do you use AI for? Where can you tell that you are learning less? Where does it actually work?" Write the points up on the board. That is your starting point.

This week: Redesign one upcoming assignment so that AI may be used along the way, but the assessment happens orally or under supervision. A concrete rewrite might look like this:

- **Before:** "Write an 800-word essay on the motif of transformation in Hans Christian Andersen's *The Little Mermaid*. Submit on Aula by Friday."
- **After:** "Prepare a five-minute oral presentation on the motif of transformation in *The Little Mermaid*. You may use AI as a thinking partner during your preparation. Bring three notes showing where the AI helped you and where you disagreed with it or corrected it. You will deliver the presentation itself without notes."

Say it out loud: "The AI may help you think. It must not write your answers."

This month: Write a short note on the parent portal about how your class is working with AI (see section 9).

7. How to test whether an AI tool is guided enough

Section 5 assumes you can judge whether a given tool is actually guided. Here are four questions you can put to the tool — and to yourself — before introducing it to students:

1. **Does it give a complete solution if you push?** Ask the tool to “write an essay on *The Little Mermaid*” or “solve this equation step by step”. If it delivers a finished solution that a student could submit as their own, it is not guided enough.
2. **Does it ask questions back?** A guided tool answers “write my essay” with something along the lines of: “What is your own main point? Which paragraphs do you already have?” If it simply hands over the content without asking questions back, it is an answer machine.
3. **Can the teacher see the student’s conversation?** If the student’s conversation with the AI is not visible to the teacher and cannot be retained, you have no way of assessing what the student actually used the AI for. That is a professional blind spot.
4. **Is it accurate at the subject level?** Ask it to explain a specific mathematical or linguistic concept from your subject. If it mixes up concepts, refers to American curricula, or gives answers that are technically wrong, it is either too poorly guided — or it is built on a model that does not know Danish K–9 schools well enough.

SkoleGPT and Miira are both designed with these principles in mind, but tools evolve. Run the test before every new school year — and again when a major update arrives.

8. When you suspect cheating

Cheating with AI is typically not a disciplinary matter — it is a sign that the task design needs work. That does not mean you should not act on it. It means the action needs to be targeted.

1. **Open a short, non-accusatory conversation.** “Tell me about how you worked on this assignment.” Listen more than you speak.
2. **Ask the student to explain two or three passages aloud, on their own.** This is the most reliable way to assess ownership. The student who used AI as a thinking partner can explain. The student who copied cannot.
3. **Avoid AI detectors.** They are unreliable and disproportionately misclassify dyslexic students and students with Danish as a second language. They create more problems than they solve.
4. **Treat the pattern, not just the incident.** If three students hand in something similar, or the whole class suddenly writes markedly better, the task is poorly designed. Rework the next one instead of only sanctioning the previous one.
5. **Agree on a plan for next time.** “On the next assignment, show me your work halfway through. Then we both know where you are.” That is a working agreement, not a punishment.

See also section 5 above: in grades 7–9, always assessing without AI, oral defences and process portfolios are the prevention. Handling cheating is the reactive part.

9. Parent communication

Parents need to know three things: what you use AI for, what their child may and may not do, and what they themselves can contribute. Keep it short. Send it twice a year: at the start of the school year and after Christmas.

To grades 1–4

Dear parents,

In your child's class we currently do not use AI as a student tool. The children have no independent access to ChatGPT or similar in the lessons. I do use AI myself to prepare materials — for instance reading tasks at several levels — but that is adult work, not something the children see or operate.

If you notice that your child is already using AI at home, do let me know. It is a useful conversation to have. We are not opposed to the technology; we simply want to make sure that the foundational skills — reading, writing, maths — fall into place before AI takes on any part of the work.

Please ask if you are in any doubt.

To grades 5–6

Dear parents,

In your child's class we are working with artificial intelligence as a thinking partner. The tool we use is called [SkoleGPT/Miira], and it has been built with GDPR safeguards. The students use it together with the teacher — not as a standalone tool for completing assignments.

This means that homework should, as a general rule, be possible to do without ChatGPT or similar. If you notice that your child is using AI in a way that gives you pause, do contact me — we will find an arrangement that works for you.

We are not opposed to the technology. We are focused on your child actually learning something. The research indicates that AI without guidance can harm learning, while AI in a good pedagogical framework can help. That is the balance we are working towards.

Please ask if you are in any doubt.

To grades 7–9

Dear parents,

Your child is welcome to use AI in connection with schoolwork — as a thinking partner, not as someone who writes the assignments for them.

In practice, this is what it looks like: Your child works at home on, say, a history assignment or a Danish writing task. They may use ChatGPT or similar to get ideas, try out counter-arguments, find phrasing, or explore the source material. But what actually gets assessed happens in class — typically as an oral presentation or as a written test in lesson time, where AI is not available. In this way, the AI may help your child *think*, but it must not deliver the work for them. The students know what the rule is, and we discuss it openly in class.

If you notice that your child is using AI to avoid the actual work — for example by simply pasting in a finished answer to a piece of homework — that is a useful conversation to have at home. We also raise it in class. If you have specific questions, do contact me.

Adapt the details to your class — tool name, rules, contact details.

10. For school leadership and municipality

A short digest. The full version is in section 09 of the research overview.

- **Run STIL's seven-step risk assessment** for each AI tool *before* it is used in the classroom (mapping, competence development, risk assessment, clarification of data responsibility, consent, oversight, documentation). STIL is the Danish Agency for IT and Learning, which runs the national school IT infrastructure.
- **Treat AI tools that effectively make assessments as high-risk** under Annex III of the EU AI Act. This covers tools that grade, place a student at a level, or whose output feeds directly into the teacher's assessment. Plain formative feedback with no consequence for official assessments typically sits below the threshold — but the line is blurry, and when in doubt, treat it as high-risk. High-risk status means a quality management system, post-deployment monitoring, and the option of human intervention in the AI's decisions. The deadline for full compliance closes in 2026–27.
- **Make AI literacy mandatory professional development for teachers.** RAND data show that training is the single thing that most predicts whether the use becomes effective.
- **Budget for training time, not just licences.** EEF and RAND consistently show that *tool plus professional development* matters far more than tool choice alone. If you must choose, choose professional development — SkoleGPT is free.
- **Avoid emotion-recognition tools entirely.** They are banned under the EU AI Act.
- **Prefer public open-source alternatives**, designed to be GDPR-secure from the ground up, such as SkoleGPT, over commercial chatbots, wherever possible. It protects you legally and pedagogically.

11. Caveats to know before citing this onwards

I repeat the most important ones from section 10 of the research overview — because you should not pass this document on without knowing its limits.

1. **The evidence base for K–9 is genuinely thin.** The strongest studies cluster around ages 13–17 or at university level. Almost no rigorous RCTs exist for grades 1–6. Assuming that what works for lower secondary also holds for early primary is professionally risky — and, as noted in section 5, the recommendations for grades 5–6 are exactly that: a deliberate but measured extrapolation.
2. **Most positive studies are short-term** (eight weeks or less). Long-term retention and lasting skill formation are largely untested.
3. **The effect sizes in meta-analyses are inflated** by publication bias (positive results are published more often than null results) and short post-tests. The moderate-to-large effects of 0.46 to 0.87 commonly cited for ChatGPT are not directly comparable to the more conservative 0.27 from the rigorous American K-12 meta-analysis of 2025. An effect size of 0.2 is small, 0.5 is moderate, 0.8 is large — but in educational contexts with uneven comparison groups, the numbers are always more indicative than definitive.
4. **The tool versions move faster than the evidence.** Most published RCTs use GPT-3.5 or GPT-4 (2023–24). The newest, most advanced AI models behave differently. Replicate before you deploy.
5. **EVA is being closed by the 2026 finance act.** That removes the most authoritative independent Danish education evaluator, precisely when the evidence is most needed in school policy.

The Nordic evidence consists mainly of surveys and qualitative studies. Use Skolverket (Swedish school authority), EVA, Udir (Norwegian Directorate for Education and Training) and the Tekstognode/Kantar-Nordic survey (a joint Nordic teacher survey from 2025) when you want to know about uptake and attitudes. Use the international RCTs when you want to know about effects. Bear in mind that the connection between the two is uncertain.

End of report. Appendices follow.

Appendix A — Reading guide to the research overview

If you want to go deeper into an area, start here:

Question	Section in the research overview
Does AI work for student learning?	01 — Effects on learning outcomes
What about motivation and special education?	02 — Motivation, engagement and special education
What do teachers get from AI?	03 — Effects on teachers
Does AI harm student thinking?	04 — Cognitive risks
What are other countries doing?	05 — Policy frameworks in the Nordics and the West
How do we get started?	06 — Implementation and competence development
Which AI tool should we choose?	07 — Comparison of AI tool types
What are the sources?	08 — Key studies
What should I do at my school?	09 — Recommendations
How strong is the evidence?	10 — Caveats

Appendix B — Glossary

- **RCT** — randomised controlled trial. Students (or teachers) are assigned at random to a group using AI and one that does not. The gold standard in educational research.
- **Pre-registered study** — a study where researchers publish in advance what they will measure and how. This prevents them from later “selecting” the positive findings. A strong quality marker.
- **K-12 / K-9** — American notation for grade levels. K-12 covers compulsory school plus upper secondary (roughly grades 0–12). K-9 in this document refers to the Danish Folkeskole (grades 0–9, ages 6–15).
- **Folkeskole** — the Danish compulsory school covering grades 0 through 9 (ages 6–15). The equivalent of UK primary and lower secondary combined.
- **STIL** — the Danish Agency for IT and Learning. Runs the national school IT infrastructure and publishes guidance on lawful AI use.
- **EVA** — the Danish Evaluation Institute. Independent evaluation body, being closed on 1 July 2026.
- **EEF** — Education Endowment Foundation. UK independent foundation funding educational trials.
- **NFER** — National Foundation for Educational Research. UK research institute.
- **RAND** — American research institute, among the most cited on teacher panel data.
- **PNAS** — *Proceedings of the National Academy of Sciences*. Highly ranked American scientific journal.
- **BJET** — *British Journal of Educational Technology*. Peer-reviewed journal on educational technology.
- **Udir** — Utdanningsdirektoratet, the Norwegian Directorate for Education and Training.
- **Skolverket** — Swedish national school authority.
- **Spørringen** — the annual Norwegian national school survey of school leaders.
- **FP9** — the Danish 9th-grade exit exams (Folkeskolens prøver, grade 9).
- **GDPR** — the EU General Data Protection Regulation of 2018.
- **Aila** — an AI lesson planner developed by Oak National Academy (UK), funded by the Department for Education. Subject of the first major K–9 RCT.
- **Generation AI** — a Finnish research project on AI literacy for students in grades 4 and 7, run by the universities of Helsinki, Oulu and Eastern Finland.
- **Aarhus CCT** — Aarhus University’s *Center for Computational Thinking and Design*. The leading Nordic research environment on AI in K–9 schools.
- **Effect size (Hedges’ g, Cohen’s d)** — statistical measure of the difference between an AI group and a control group. 0.2 small, 0.5 moderate, 0.8 large.
- **Publication bias** — the tendency for studies with positive results to be published more often than null findings. This skews meta-analyses upwards.
- **Scaffolding** — pedagogical concept where the teacher (or the tool) supports the student gradually, allowing them to take steps that would otherwise be too large.

-
- **Socratic question** — a question that does not reveal the answer but leads the student to think further. “What happens if you try with a smaller number?” rather than “the answer is 4”.
 - **Process portfolio** — collected documentation of the student’s working process (drafts, notes, reflections, not just the final product). Well-suited to making AI use visible.
 - **Hallucination** — when an AI model produces something that sounds correct but is wrong or fabricated.
 - **Annex III of the EU AI Act** — the list of high-risk AI uses, which includes AI used for assessment in education.
 - **ISAE 3000 Type II** — international auditing standard for data processors; demonstrates that a vendor’s data processing has been independently audited over an extended period.
 - **Aula** — the Danish school portal used for parent–school communication and homework distribution.
-
-

This document is a digest. The full research overview, including references and effect sizes, is in “Artificial Intelligence in Danish K–9 Schools — research overview, May 2026” by Martin Hasbirk Brandt.