The Sethi Ledger

PEOPLE / PRACTICE / POSSIBILITY

LEADERSHIP • WEEKEND RESEARCH SUPPLEMENT

Keep thinking visible when AI helps

October 17-23, 2026

This week's research asks what learners and teachers do before, during and after an AI response. School studies make model-building and teacher handoffs tangible. University evidence offers cautious ideas for questioning, feedback and planning; the consistent leadership task is to examine the thinking behind a polished product.

The week at a glance

01 Build explicit routes from an AI tutor to a teacher.

02 Make student reasoning visible before and after AI assistance.

03 Distinguish reported confidence, supported performance and independent learning.

04 Fund review and evaluation alongside access to the tools.

ABOUT THIS BRIEF

Selected research published October 17-23, 2026. Findings are assessed for relevance to K-12 leadership, evidence quality and practical implications. Each summary identifies important limitations.

01 / LEADERSHIP

Make the handoff to a teacher explicit

Peer-reviewed exploratory design pilot
Payevsky and colleagues / Education Sciences

THE FINDING

When should an AI tutor hand a struggling learner back to a teacher? A six-week algebra pilot with 32 eighth-grade boys combined a rule-based practice system, a generative AI questioning partner and direct teacher support. Classroom use exposed practical weaknesses: an English interface burdened Hebrew-speaking learners, and open-ended questions alone sometimes left students stuck. The team developed a local-language practice tool, added more explicit guidance and introduced a way to call the teacher. These are useful implementation findings about coordinating different kinds of help. The pilot did not test whether the arrangement improved learning, and its proposed redistribution of cognitive load remains a theory to evaluate.

THE EVIDENCE

Observations and feedback refined seven instructional units on systems of equations. The teacher also conducted the research; the practice tool used predefined response rules.

LIMITS TO KEEP IN VIEW

One high-achieving class in a religious boys' school; no effectiveness comparison. Feasibility does not establish gains in achievement or reduced cognitive load.

WHAT LEADERS CAN DO

Require every tutoring pilot to specify signs of confusion, the escalation route and an adult response time. Test the interface in pupils' instructional language. Evaluate unaided mathematics separately from successful completion with support.

WHO / WHAT WAS STUDIED

Grade 8, Haifa, Israel; direct school evidence.

READ THE ORIGINAL SOURCES

Human-AI orchestration: full research article
mdpi-res.com · Published 2026-10-22 · Open source

02 / LEADERSHIP

Let pupils build and question a model

Peer-reviewed pre/post classroom study
Redondo Duarte and Saez Lopez / IJoSI

THE FINDING

Can building a simple model help children understand AI? Across 11 Madrid-area schools, 996 pupils aged 10-11 completed four sessions using visual programming and tools such as Teachable Machine. They created and trained models rather than only consuming AI-generated answers. Average scores on a ten-point test rose from 5.54 to 6.24; the test covered machine-learning foundations and supporting programming concepts. Nearly 95% reported enjoying model creation. This is encouraging evidence that an accessible, practical unit can accompany improved knowledge and strong interest. It cannot show that the unit caused the gains: there was no comparison group, and the analysis did not account for pupils sharing classrooms and schools.

THE EVIDENCE

Purposive sample from 17 classrooms; paired pre/post tests reported a standardized change of about 0.51. The article identifies this dataset as distinct from earlier related work.

LIMITS TO KEEP IN VIEW

Immediate results, unmeasured prior programming experience and clustered pupils limit inference. Enjoyment and knowledge scores do not establish broader critical-thinking transfer.

WHAT LEADERS CAN DO

Pilot a short build-test-explain unit. Ask pupils to change training examples, find a failure and explain why it occurred. Include an unfamiliar classification task afterward, with individual explanations and accessible alternatives for pupils who need them.

WHO / WHAT WAS STUDIED

Primary pupils in Spain; direct K-12 evidence.

READ THE ORIGINAL SOURCES

Creating artificial intelligence: full research article
article.accscience.com · Published 2026-10-22 · Open source

03 / LEADERSHIP

Support teachers before expanding AI use

Peer-reviewed mixed-methods study
Yafie and colleagues / Education Sciences

THE FINDING

What helps kindergarten teachers use AI responsibly? Researchers in Malang, Indonesia, interviewed ten teacher leaders and five administrators, then surveyed 250 kindergarten teachers with prior AI-use experience. Teachers described planning stories, activities, media and family communications while checking developmental fit, culture, accuracy and safety. Perceived usefulness, ease of use and enabling conditions were associated with intention to use AI; social influence was not statistically significant. Infrastructure, training, premium-feature restrictions and privacy remained obstacles. The study points to the practical support surrounding adoption, not a mandate to increase children's screen time. Because non-users were excluded, its results describe experienced users rather than the readiness of the entire teaching workforce.

THE EVIDENCE

Cross-sectional modeling used equal samples from five districts. Findings concern teachers' reported professional practices; the study did not measure child learning outcomes.

LIMITS TO KEEP IN VIEW

One city, self-reported use and unweighted district allocation. Associations cannot establish that training or institutional support caused greater integration.

WHAT LEADERS CAN DO

Offer protected planning time and practice reviewing generated materials for age, language and cultural fit. Include non-users in a local needs assessment. Price the full workflow, including access, review time and privacy support, before choosing a license tier.

WHO / WHAT WAS STUDIED

Kindergarten educators in Indonesia; direct workforce evidence, not a child trial.

READ THE ORIGINAL SOURCES

Kindergarten AI integration: full research article
mdpi-res.com · Published 2026-10-21 · Open source

04 / LEADERSHIP

Teach a sequence of questions

Peer-reviewed quasi-experimental study
Lin and colleagues / IJETHE

THE FINDING

Does a structured questioning routine improve how learners work with AI? In a television-production course, 83 undergraduates in three intact classes used DeepSeek for a short-film project. One class asked free-form questions; two followed progressive or divergent question chains. Both structured groups showed more strategy and monitoring behaviors and higher self-reported metacognitive awareness. Mean course scores were 90.50 and 90.13, compared with 86.67 in the free-form class; the two structured approaches did not differ significantly. The finding supports examining the whole instructional routine. It does not isolate the value of chaining: the groups also received different question specificity, completion requirements and reflection prompts, and individual baseline knowledge was not measured.

THE EVIDENCE

A five-week project followed common foundational instruction. Researchers analyzed interaction logs, reflections, post-task awareness measures and course grades.

LIMITS TO KEEP IN VIEW

Three classes at one university; no individual randomization. Scaffolds stayed available, so independent transfer was not tested. Reflection prompts could cue reported themes.

WHAT LEADERS CAN DO

Have teachers develop a brief routine that asks for alternatives, evidence and revision. Compare work using the routine with a matched task, then remove the prompts to check what students can regulate independently.

WHO / WHAT WAS STUDIED

Chinese university students; indirect evidence for K-12 instruction.

READ THE ORIGINAL SOURCES

Question-chain prompting: full research article
link.springer.com · Published 2026-10-23 · Open source

05 / LEADERSHIP

Make feedback a judgment task

Peer-reviewed accepted article / quasi-experiment
Yan and Meng / HSS Communications

THE FINDING

Receiving comments is different from learning to evaluate and use them. A ten-week study assigned three intact university English classes to AI, peer or teacher feedback, with 126 students completing the analysis. AI and peer groups improved their reported writing-feedback literacy and scored above the teacher group after adjustment for baseline scores; AI and peer groups did not differ significantly. Interviews suggested that students engaged more actively in interpreting those two sources. This concerns how learners handle feedback, rather than proof of better writing or replacement of teachers. The same instructor taught all classes, and each feedback condition corresponded to a single class, leaving classroom differences difficult to disentangle.

THE EVIDENCE

The 28-item literacy scale measured perceptions across five dimensions. The adjusted group effect was partial eta-squared .067; interviews added process accounts.

LIMITS TO KEEP IN VIEW

One university, short follow-up and self-report outcomes. Class and treatment are intertwined. This accepted early version remains subject to publisher editing.

WHAT LEADERS CAN DO

Ask learners to compare two suggestions, accept or reject each with reasons, and show the resulting revision. Keep the teacher responsible for feedback quality. Score the revised writing separately from confidence or satisfaction with the feedback process.

WHO / WHAT WAS STUDIED

University English learners aged 17-20 in China; indirect K-12 evidence.

READ THE ORIGINAL SOURCES

AI, peer and teacher feedback: accepted manuscript
www.nature.com · Published 2026-10-23 · Open source

06 / LEADERSHIP

Put the argument on paper first

Peer-reviewed qualitative design research
Collins / Contemporary Educational Technology

THE FINDING

A paper-first routine asks students to map their argument before consulting AI. Over two years and nine design cycles, a South African university project developed this approach with a six-person design team and a cohort of 150 foundation-level students. Learners drew logic diagrams before using digital tools, then used those diagrams to check the resulting text. Observations and focus groups described clearer reasoning, fewer trial-and-error prompts and better recognition of weaknesses in an argument. Some students initially resisted the extra work. These findings offer a concrete way to preserve authorship and make reasoning visible. They are qualitative design evidence, not an estimate of learning gains or proof that paper itself is essential.

THE EVIDENCE

Four student focus groups and ten design-team sessions informed iterative refinement. More than 85% of the cohort used English as an additional language.

LIMITS TO KEEP IN VIEW

One institution, purposive participation and no comparison condition. Reported efficiency and ownership were not independently established through standardized outcome tests.

WHAT LEADERS CAN DO

Try a short argument map before an AI-assisted writing task. Allow an accessible equivalent where handwriting is a barrier. Ask students to identify where the generated draft departed from their plan and justify the final reasoning in their own words.

WHO / WHAT WAS STUDIED

First-year ICT students, South Africa; indirect evidence for school writing.

READ THE ORIGINAL SOURCES

Paper-first knowledge engineering: full research article
www.cedtech.net · Published 2026-10-19 · Open source

07 / LEADERSHIP

Measure what human oversight actually does

Peer-reviewed qualitative systematic review
Spyropoulou and Kameas / Computers

THE FINDING

Human oversight is often promised more clearly than it is evaluated. A review of 78 higher-education studies examined educators' use of generative AI to design instruction, the competencies involved and professional learning approaches. Design-stage activity appeared in 95% of studies, while implementation and evaluation received much less attention. Oversight was explicitly endorsed in 58%, but empirically examined in only 21 studies, or 27%. Most studies used no named competence framework. The practical message is that generating a lesson and checking whether it works are different professional tasks. The review supports closer examination of teachers' review behaviors; it does not provide a pooled estimate that a particular training model improves student achievement.

THE EVIDENCE

PRISMA-guided searches covered 2022-2026. The synthesis included practicing educators and preservice contexts and used methodological appraisal to qualify confidence.

LIMITS TO KEEP IN VIEW

English-only searches, heterogeneous studies, substantial self-report and single-reviewer screening/coding. Twenty-two reports could not be retrieved; quantitative pooling was inappropriate.

WHAT LEADERS CAN DO

Make professional learning produce evidence of review: a corrected misconception, an identified bias and a justified instructional revision. Follow the resulting lesson into classroom use. Allocate evaluation time alongside content-generation time when budgeting for adoption.

WHO / WHAT WAS STUDIED

Higher education and teacher preparation; indirect evidence for district professional learning.

READ THE ORIGINAL SOURCES

GenAI instructional design: full systematic review
mdpi-res.com · Published 2026-10-22 · Open source

08 / LEADERSHIP

Turn AI policy into usable guidance

Peer-reviewed mixed-methods study
Suing and colleagues / Frontiers in Education

THE FINDING

Students can value AI while still wanting clearer rules. Research in Ecuador combined interviews with 14 lecturers and eight students, a 200-student survey and institutional document review. Participants described useful assistance with academic work alongside difficulty verifying outputs, uneven access and concerns about dependence. Survey respondents rated demand for regulation at 4.00 on a five-point scale, compared with 3.21 for perceived clarity of existing rules. The gap suggests a practical governance problem: a policy document may exist without making everyday decisions understandable. These results characterize the participating university sample. They do not establish national prevalence, prove that AI weakens critical thinking, or show that stricter rules improve learning.

THE EVIDENCE

The survey used online snowball recruitment. Interviews and policy review supplied context for students' perceptions and reported practices.

LIMITS TO KEEP IN VIEW

Nonrepresentative adult sample and self-report. The two-item governance construct had poor reliability (alpha .203); its demand and clarity items need separate interpretation.

WHAT LEADERS CAN DO

Test district guidance with realistic classroom scenarios: permitted help, source checking, disclosure and a disputed answer. Ask teachers and students to explain the expected response. Revise unclear examples before treating publication of the policy as implementation.

WHO / WHAT WAS STUDIED

Social-science university communities in Ecuador; indirect K-12 governance evidence.

READ THE ORIGINAL SOURCES

University AI practices and governance in Ecuador
www.frontiersin.org · Published 2026-10-19 · Open source

FROM EVIDENCE TO ACTION

The week ahead

01 Instruction: ask one school team to test a short explain-before-AI routine and an unfamiliar unaided task afterward.

02 Student support: define when a tutoring system must hand a problem to a teacher, including language and accessibility needs.

03 Professional learning: use real generated materials to practice spotting factual errors, bias and unsuitable developmental demands.

04 Evidence: require pilot owners to name the measured outcome and the strongest alternative explanation for a reported gain.

05 Governance: test AI guidance with teachers and students using concrete classroom scenarios; revise the examples they cannot apply.

KEEP READING

Visit The Sethi Ledger · Explore the archive