

AI, Schools and Student Thinking

September 6-13, 2026 | Early edition through September 12. Ten new publications selected for a large school division's superintendent. September 13 releases are not yet covered; the final weekly selection may change.

Executive takeaways

Check independent learning. PISA and verification research reinforce the need to assess students' own reasoning.^{1,3,4}

Protect human connection. New adult studies flag risks around sustained companionship and reported psychological harms.^{5,6}

Pilot with clear aims. Teacher rehearsal shows promise; define what students and educators must learn before scaling.^{2,7,8,9}

Reading priorities

Click a topic for its summary. Each summary links to the original source. Ranking reflects leadership relevance and evidence quality, not citation counts.

No.	Topic and finding	Evidence / release
1	PISA: AI use and student performance How students use AI matters; associations do not establish cause.	OECD assessment September 8
2	AI rehearsal for elementary teachers Lesson-planning gains were stronger than content-knowledge gains.	Journal / comparison September 10
3	Explanations that help users find errors Preferred AI explanations were not always easiest to verify.	arXiv / experiment September 8
4	AI participants in ethical discussion Perspective-taking improved by self-report; agreement could blunt debate.	Preprint / user study September 10
5	AI companionship and human connection Sustained companionship use accompanied lower well-being.	Preprint / longitudinal September 7
6	Reported psychological harms from chatbots Serious safety signals warrant attention; prevalence remains unknown.	Preprint / harm reports September 7
7	UNESCO: public purpose and learner agency Keep educational aims central to AI decisions.	UNESCO / analysis September 8
8	Different expectations among staff and students Staff and students did not share a clear view of appropriate use.	Journal / survey September 10
9	A practical distinction between help and delegation The same AI tool can support very different learning processes.	Preprint / conceptual September 9
10	A null finding on AI use and self-esteem Frequency of use showed no detectable link to performance self-esteem.	Journal / survey September 7

Evidence mix: 3 journal studies, 5 arXiv manuscripts, and 2 institutional publications. Several studies concern adults; two entries offer expert or conceptual analysis rather than new outcome data.

Student learning and teacher preparation

1. PISA: AI use and student performance

September 8, 2026 | International assessment and survey report

Direct K-12 evidence: 15-year-olds, 91 systems

Finding. OECD reading scores fell 14 points from 2022. PISA links AI use for summarizing, drafting, and preliminary research with lower science scores. Weekly use for general learning shows similar scores to non-use; opportunities to assess AI information accompany better results among frequent users.¹

Evidence. More than 760,000 students participated across 91 countries and economies. AI analyses combine assessed performance with reported use and adjust for socioeconomic status.

Limits. These are associations. They cannot establish that AI caused score declines or that AI-literacy lessons caused improvement. Disadvantaged students report fewer opportunities to evaluate AI information.

Leadership implication. Protect independent reading and reasoning. Teach students to check AI claims, and ensure those lessons reach every school. Use local evidence of unaided mastery to evaluate pilots.

2. AI rehearsal for elementary teachers

September 10, 2026 | Peer-reviewed quasi-experimental study

K-12 workforce: teacher candidates, Kazakhstan

Finding. Teacher candidates using chatbot simulations improved more in lesson planning, pedagogical knowledge, and teaching confidence than the comparison group. The content-knowledge difference was not statistically significant.²

Evidence. Ninety-one analyzed participants: 44 received eight sessions over four weeks; 47 received no simulation intervention. Human raters scored paper-and-pencil lesson plans.

Limits. Participants chose their groups, and only the intervention group required prior AI experience. Extra practice and selection differences could explain some gains. No pupil outcomes or long-term classroom transfer were measured.

Leadership implication. Consider supervised rehearsal in teacher development. Judge success through independently assessed plans and subsequent classroom practice, including subject accuracy, before scaling.

1. OECD, PISA 2025, Chapter 4 and Box I.4.2.

2. Abukhanova et al., teacher simulations, Sections 3-5 and Table 2.

Critical thinking and ethical judgment

3. Explanations that help users find errors

September 8, 2026 | Author manuscript: controlled human study

University adults: indirect K-12 implications

Finding. Participants preferred plans and decomposed explanations, but simpler step-by-step traces often better supported verification. Some preferred formats increased false alarms or encouraged trust without corresponding willingness to check.³

Evidence. Fifty adult students completed a controlled comparison of six formats, using 27 selected problems and deliberately inserted reasoning errors. Assignment and presentation order were randomized.

Limits. One university sample, selected benchmark problems, and artificial errors limit classroom generalization. The study measures immediate evaluation, not lasting critical-thinking gains or authentic internal model reasoning.

Leadership implication. Test whether students can locate and explain a flaw in an AI answer. Treat satisfaction with an explanation as a separate measure from successful checking.

4. AI participants in ethical discussion

September 10, 2026 | Preprint: mixed-methods user study

University STEM students: indirect K-12 implications

Finding. Students reported greater ethical awareness and reasoning confidence after discussions with AI personas and peers. Interviews also identified overly agreeable AI responses and limited interaction between human participants.⁴

Evidence. Forty-five undergraduates worked in 15 groups during a roughly two-hour session. Researchers combined before-and-after self-ratings, discussion logs, peer ratings, and interviews.

Limits. There was no comparison group without AI. Reported sensitivity is not independently demonstrated ethical reasoning; a brief session cannot establish durable learning or suitability for children.

Leadership implication. For a discussion pilot, require students to challenge claims, compare viewpoints, and explain their own conclusion. Preserve time for peer disagreement and teacher-led discussion.

3. Lim et al., reasoning representations, Section 5 and Appendix A.4.

4. Seo et al., Ethics Training Agents, Sections 5-7.

Well-being and student protection

5. AI companionship and human connection

September 7, 2026 | Preprint: longitudinal observational study

U.S. adults, ages 18-90: no child sample

Finding. Sustained social engagement with Character.AI was associated with lower well-being. Less in-person interaction statistically explained much of that relationship.⁵

Evidence. Researchers surveyed 1,182 adult users initially and 439 approximately a year later. Models accounted for baseline measures and attempted to correct for selective follow-up.

Limits. Only 37% returned. Human interaction was measured only at follow-up, limiting claims about displacement or causal direction. Self-report, one platform, and an adult sample limit application to students.

Leadership implication. Review whether school-supported tools encourage reliance or extend human support. In family guidance, distinguish an instructional assistant from a companion designed for sustained social engagement.

6. Reported psychological harms from chatbots

September 7, 2026 | Preprint: analysis of reported adverse experiences

Predominantly adult reports: school safeguarding context

Finding. Clinician raters identified descriptions consistent with delusional beliefs in 102 of 185 harm reports; 50 of those described chatbot validation. Reported consequences included isolation, relationship breakdown, and hospital admission.⁶

Evidence. The dataset contains 95 first-hand and 90 second-hand accounts submitted to a support organization. Median reported age was 35; researchers independently coded deidentified accounts.

Limits. Reports were self-selected, retrospective, and unverified. These proportions describe a harm-reporting sample, not all chatbot users. The study establishes neither causation nor risk rates for children.

Leadership implication. Review how student-facing systems respond when conversations involve distress or implausible beliefs. Require clear routes to trained human support and a process for reporting concerning interactions.

5. Zhang et al., AI companionship, Sections 2-4.

6. Morrin et al., reported chatbot harms, Sections 2-4.

Public purpose and shared expectations

7. UNESCO: public purpose and learner agency

September 8, 2026 | Institutional collection of expert analysis

Education-system governance and child development

Finding. UNESCO frames AI adoption around student development, autonomy, pedagogy, assessment, expression, and public accountability. It calls on education systems to actively shape their response to algorithmic technologies.⁷

Evidence. The volume assembles 26 think pieces around six questions about education's future, including who retains oversight and how young people develop cognitively, emotionally, and socially.

Limits. This is expert analysis, not an intervention evaluation or pooled estimate of effects. The publisher's introduction was reviewed; the full collection could not be accessed for this brief.

Leadership implication. Use the questions in a cabinet discussion: what should remain under educator control, what thinking must students do themselves, and how will families hold the division accountable?

8. Different expectations among staff and students

September 10, 2026 | Peer-reviewed cross-sectional survey

One UK university: indirect K-12 implications

Finding. Students were more positive about AI than staff, while staff expertise and perceived uses varied. Student uses centered on language improvement and text production; appropriate academic use remained ambiguous.⁸

Evidence. Surveys collected in April-June 2024 included 86 complete student responses and 45 complete academic responses, with additional partial responses and two teaching assistants. Most responding students were postgraduates.

Limits. Low participation and self-selection limit representativeness. These are perceptions from 2024, not current adoption rates or measured learning outcomes. Different performance-expectation scales constrain direct staff-student comparisons.

Leadership implication. Discuss concrete assignments with teachers and students to establish shared expectations. Recheck local understanding as tools change, rather than relying on a one-time AI orientation.

7. UNESCO, *The algorithm in the room*, publisher introduction.

8. Kahn et al., *staff/student survey*, Methodology and Findings.

Learning processes and psychological effects

9. A practical distinction between help and delegation

September 9, 2026 | Preprint: conceptual education analysis

Instructional design across educational settings

Finding. A smart-kitchen-appliance analogy illustrates how AI can automate an outcome or help a learner practice. The authors argue that the mode of use and resulting cognitive activity determine its educational value.⁹

Evidence. The paper maps illustrative AI uses onto established frameworks for cognitive engagement and technology integration. It presents a conceptual analysis, with no new student experiment.

Limits. An analogy explains distinctions but does not test them. It cannot establish that AI generally causes skill loss or that any particular classroom design improves learning.

Leadership implication. For each proposed use, ask teachers to identify the mental work students retain: generating an explanation, evaluating alternatives, revising reasoning, or applying knowledge independently.

10. A null finding on AI use and self-esteem

September 7, 2026 | Student-led journal study: adult survey

Adult convenience sample: no adolescent outcomes

Finding. The hypothesized relationship between frequent AI use and lower performance self-esteem was not supported. The observed correlation was near zero ($r = -0.039$; $p = 0.675$).¹⁰

Evidence. A survey of 121 adults assessed AI-use frequency, upward social comparison, and performance-related self-esteem. Researchers analyzed correlations rather than assigning AI use.

Limits. Convenience recruitment, a single frequency question, and no non-user group constrain interpretation. A non-significant association does not establish safety; adolescent effects and changes over time were not measured.

Leadership implication. Keep risk claims specific to the outcome and population studied. Include student confidence and well-being in local evaluation without presuming that every type of AI use has the same psychological effect.

9. Rummel et al., kitchen-machine analogy, arXiv:2609.09856v1.

10. Kim and Cho, AI use and self-esteem, results and limitations.

Sources and selection notes

Selection and interpretation

Requested window: September 6-13, 2026, inclusive; search cutoff: September 12 (UTC). The window overlaps the prior edition on September 6. Dates use first-online publication, initial arXiv posting, or an institutional release. This early selection excludes future-dated releases and new coverage of older studies. Ranking weighs leadership relevance, evidence strength, K-12 directness, and distinct findings. It is a curated selection, not a systematic review.

Leadership implications are analytical judgments. Adult findings require caution in K-12 use; publication status does not establish causation. UNESCO is summarized from its publisher introduction; full text or relevant report sections were reviewed for the other entries.

1. OECD. [PISA 2025 Results \(Volume I\): Future-Ready Students](#). OECD Publishing. September 8, 2026. Open report; Chapter 4, pp. 236-241, especially Box I.4.2. [Open AI section](#).

2. Azhar Abukhanova et al. [Towards professionalism: AI to benefit student-teachers' lesson planning](#). *Frontiers in Education*; DOI 10.3389/feduc.2026.1833547. September 10, 2026. Open full text; methods in Section 3, outcomes in Table 2.

3. Jaewoo Lim, Sungbok Shin, and Sanghyun Hong. [Do Reasoning Representations Help Humans Evaluate LLM Outputs?](#). arXiv:2609.09038v1. September 8, 2026. Open full text; arXiv record lists EMNLP 2026 Findings. Initial posting date used.

4. Youngseok Seo et al. [Ethics Training Agents: Facilitating Group-Based Ethics Education with Role-Playing and Discussion for Ethical Reflection and Exploration](#). arXiv:2609.11529v1. September 10, 2026. Open full text; methods in Section 5, limitations in Section 7.4.

5. Yutong Zhang et al. [Living with AI Companions: Sustained AI Companionship Predicts Lower Well-Being Through Lower Human Interaction](#). arXiv:2609.07243v1. September 7, 2026. Open full text; sampling and attrition in Sections 3.1 and 4.

6. Hamilton Morrin et al. [Delusions and Harms Associated With AI Chatbot Use: Early Evidence From 185 Real-World Reports](#). arXiv:2609.08027v1. September 7, 2026. Open full text; methods, denominators, and limitations in Sections 2-4.

7. UNESCO. [The algorithm in the room: Seeing and confronting the implications of AI for the future of education](#). UNESCO. September 8, 2026. Publisher introduction reviewed; full collection linked from that page. First release September 8; page updated September 9.

8. Peter Kahn et al. [Still emerging: understanding Generative AI use in Higher Education](#). *Frontiers in Education*; DOI 10.3389/feduc.2026.1885253. September 10, 2026. Open full text; see Data collection and Findings. Publication date differs from data-collection year.

9. Nikol Rummel, Valentina Nachtigall, and Ernesto Panadero. [With a Thermomix You Lose the Ability to Cook: A Kitchen Machine Analogy for Applications of Generative AI in Education](#). arXiv:2609.09856v1. September 9, 2026. Open 26-page PDF; conceptual paper, not an experimental finding.

10. Jayden Minjoon Kim and Jung Yeun Cho. [It's better than me: Investigating the psychological risks of AI on self-esteem](#). *Journal of Emerging Investigators*; DOI 10.59720/25-237. September 7, 2026. Open full PDF; primary correlation on printed p. 3, limitations on pp. 5-6.

Selection exclusions

PISA news coverage and country notes are not counted as separate studies. The September issue of [Online Learning](#) was published September 1, before this window. Conceptual papers are labeled explicitly; they should not be read as evidence of measured learning gains.