

A NEW DIRECTION FOR STUDENTS IN AN AI WORLD: PROSPER, PREPARE, PROTECT

MARY BURNS, REBECCA WINTHROP, NATASHA LUTHER,
EMMA VENETIS AND RIDA KARIM



Benefit 3: AI can improve student learning

AI explains things better.

— Parent, student, teacher

AI provides a wide range of potential and actual benefits to students' learning and development through both direct and indirect pathways. As is often the case with new technologies, there is a lag between the use of AI and rigorous research documenting the accumulated benefits on student learning. This reality notwithstanding, interviews with study participants suggest that AI applications, when implemented in pedagogically and developmentally appropriate ways, can meaningfully enhance student learning outcomes across multiple domains. Research by and with education and technology companies also shows that embedding AI features—such as interactive question-and-answer opportunities—within textbooks and other static, text-based teaching and learning materials improves student learning outcomes across a range of subjects, from biology to history (Anderson 2025; LearnLM Team and Google 2025; Microsoft 2025a; Pearson 2025). In our study, 62% of teacher responses suggest that AI can improve student learning by empowering students. When asked which elements of learning will gain most from AI, 35% of teacher responses point to the two domains most traditionally associated with education: reading and writing, explored below.

SUPPORTING STUDENT READING, ESPECIALLY FOR SECOND LANGUAGE LEARNERS

Long before the widespread adoption of AI, technology-enhanced learning has long provided a more engaging and effective approach to reading instruction and language instruction compared to traditional methods (Burns 2021). Teachers

interviewed for this study enumerated several elements of AI that lead to improved student reading comprehension—particularly for those learning a second language (L2 learners). Many of these advantages directly address the linguistic foundations of reading, including vocabulary knowledge and syntactic understanding. Examples include the capacity of AI to simplify vocabulary, modify reading comprehension questions, and adjust the Lexile (or reading difficulty) levels of texts to match individual learners' proficiency levels appropriately.

However, the benefits of AI also extend to important nonlinguistic factors that, while unrelated to language per se, remain crucial for reading comprehension. Teachers attest to the ability of AI to engage students in the reading process, scaffold learning strategies, extend reading time as needed, utilize text-to-speech features, and create zones of privacy and comfort where student reading difficulties remain confidential.

While research on AI and reading comprehension is scant at present, with studies of varying quality, the evidence that does exist supports teacher assertions regarding the capacity of AI to enhance reading comprehension for L2 learners. A range of studies, employing different methodologies, from a spectrum of countries (Kuwait, Saudi Arabia, India, and the Philippines), student age cohorts, and a range of AI tools (ITSs, personalized learning platforms, and general-purpose AI platforms) corroborate teacher claims (Alazemi 2024; Ali et al. 2023; Wangdi and Shimray 2025).

For example, a study from Kuwait appears to show promising results for AI's impact on reading comprehension and language-learning outcomes. Eighty grade 9 female English-language learners were selected through convenience sampling and divided equally into treatment and control groups comprising 40 students each. Researchers administered pretests that were conducted to assess their initial levels of English-language comprehension. The treatment group received AI-based instruction through a commercial

multimedia platform, *Nearpod*, while the control group received traditional face-to-face instruction. Both groups were taught 10 reading passages, and post-tests were administered to measure reading comprehension progress. While both groups exhibited parity in their pretest reading scores, the post-test showed a small improvement in the intervention group, which scored 17.75 compared to 16 for the control group. Statistical analysis confirmed that this 1.75-point difference was unlikely to have occurred by chance, indicating a genuine if modest effect (Alazemi 2024).

The above findings, along with information provided through teacher interviews, suggest that AI-based personalized reading platforms may be particularly effective in improving reading comprehension, especially for L2 learners. Much of this success appears linked to the ability of AI to match the level of reading difficulty to individual student reading abilities, analyze vast amounts of data to provide targeted insights, feedback, and personalized support to students, as well as to engage students in more individualized ways (such as asking more personalized questions), thus enhancing student engagement and motivation. Personalization will be discussed in the next benefit.

HELPING STUDENTS WITH THE WRITING PROCESS

AI holds significant promise for enhancing writing instruction and student outcomes across multiple dimensions of the writing process when used to amplify student learning rather than shortcut it. Interviews with teachers and education experts, alongside emerging research, suggest that AI technologies can support both teachers and students across the entire writing continuum—when thoughtfully integrated into pedagogical frameworks. This integration requires balancing structured writing with more exploratory and iterative practices and maintaining equilibrium between human guidance and technological support (Ferman et al. 2025; Deng et al. 2025; Guo and Wang 2024; Dell’Acqua et al. 2023).

The use of AI tools to support student writing offers three benefits:

Process-based writing support: When used as part of a pedagogical approach that supports (not replaces) student writing, AI technologies can provide meaningful assistance in the writing process. Teachers in our study report that students are more motivated to write when using AI as a scaffolding tool. This effective and motivational support of AI can be particularly valuable for students who have writing anxiety, who are weak writers, or who have difficulty conceptualizing writing topics or formulating written thoughts. It can also supplement the traditional writing process in schools which often focuses on assigning, rather than teaching, writing.

AI can help students in all phases of the writing process. During the ideation and conceptualization phase, teachers report that AI can “spark creativity” and help students overcome writer’s block by helping them “get unstuck” so they can develop arguments and ideas. At the drafting stage, it can help with organization, coherence, syntax, semantics, and grammar. At the revision stage, AI can support the editing and rewriting of ideas as well as help with the mechanics of writing, including punctuation, capitalization, and grammar.

Feedback: In *Visible Learning*, his groundbreaking meta-analysis of over 50,000 education studies, John Hattie highlighted the profound impact of feedback on student learning outcomes, ranking it among the most powerful educational interventions in terms of average effect size (2010). Feedback plays a particularly important role in supporting students’ writing revision skills because it operates on both an academic and an affective level, helping students advance their writing skills while cultivating the confidence needed to complete writing tasks (Graham 2019; Meyer et al. 2024). Yet high-quality feedback is time consuming, and teachers can only support and guide a certain number of students at any one time. To counter this constraint, teachers are increasingly using technologies like LLMs, as well Automated Written

Evaluation (AWE) systems, such as Grammarly, to automate assessment and feedback. AWE systems typically comprise two essential components: a scoring engine that produces automatic scores, and a feedback engine that delivers Automated Written Corrective Feedback.

There is growing evidence that this array of AI tools can be used to identify individual writing needs and provide meaningful, personalized feedback to improve both the process and product of writing. Meyer et al. (2024) randomly assigned 200 German secondary school students to receive LLM feedback after writing an essay draft in English (the treatment group) while 250 students received no LLM feedback on their essay (the control group). Those receiving LLM feedback experienced very modest statistically significant improvements in text revision ($d = 0.19$). However, more notable benefits occurred in how students felt about writing (motivation and emotions) with medium improvements in task motivation ($d = 0.36$), and positive emotions ($d = 0.34$).

Though these impacts are quite modest, the researchers advocate for the use of AI to create feedback opportunities for students that would be otherwise unavailable given teacher workload: “Even if the beneficial effect of one single LLM-generated feedback session is relatively small for revision performance, the effects of further feedback could cumulate for students because the automation allows for multiple feedback cycles and opportunities for practice” (Meyer et al. 2024, 7).

While AI demonstrates strengths in writing support, research also emphasizes the complementary nature of AI and human feedback. The University of California at Davis’s curricular intervention, Peer & AI Review + Reflection (PAIRR), blends both AI and human feedback to support students during the writing process, following up with critical reflection on the process and its impact on their learning (University of California at Davis 2025). One study of 400 feedback instances for the same essays (200 pieces of human-generated formative feedback and 200 pieces

of AI-generated formative feedback) found that human raters provided superior feedback in terms of clarity, accuracy, supportive tone, and emphasis on critical improvement areas, while AI feedback excelled in delivering criteria-based evaluations (Steiss et al. 2024). This research suggests that optimal writing instruction may benefit from hybrid vigor, leveraging and integrating both AI and human feedback with AI (in formative early drafts or instances where a well-trained teacher is unavailable) but relying on human teachers for more advanced stages of the writing process.

Improved writing outcomes: A number of quantitative studies (Meyer et al. 2024; Guo and Wang 2024) and qualitative ones (Meniado et al. 2024; Wang 2025) point to the role of AI, when used as “an active collaborator,” to create demonstrable improvements in student writing (Meniado et al. 2024, 7). As with AI tools for reading, these benefits are especially pronounced for L2 learners and involve the production of coherent texts with clear ideas, improved writing flow, and discourse maturity that approaches native-speaker-like quality, with significantly higher indexes of lexical diversity reflected in both ideation and sentence fluency. AI does this by accelerating the writing process, easing cognitive load, fostering new learning opportunities, providing immediate feedback, and promoting positive feelings about writing (Wang 2025).

While AI for writing studies examined in this study focuses on LLMs, AI-based writing tools have become increasingly popular writing supports. Kaliisa et al.’s (2025) meta-analysis combined findings from 41 separate studies involving nearly 5,000 secondary and post-secondary students to examine the effectiveness of AI feedback on student writing. The research determined that students who received AI feedback learned just as much as those who received traditional teacher feedback, suggesting that AI feedback is essentially equivalent to human feedback in terms of measurable learning outcomes. Rather than advocating for one approach over the other, the researchers recommend that educators use both

strategically. This hybrid model leverages what AI does well—providing quick, consistent feedback at scale—while preserving what human teachers do best—offering empathetic, nuanced, and deeply contextual responses.

An important caveat in terms of examining the effects of AI to help student writers focuses on the type of AI used. Many studies examining AI tools in education conflate fundamentally different types of AI systems, such as rule-based approaches, machine learning models, and generative AI. This conflation is problematic because these technologies operate on distinct principles and offer qualitatively different capabilities.¹ Treating these diverse technologies as equivalent when evaluating their educational impact obscures important distinctions in how they function and what they can reasonably accomplish in supporting student learning.

CONCLUSION

Given the prevalence of reading and writing difficulties among students, especially L2 learners, the above research findings hold important implications for teachers seeking to enhance students' reading skills. The unique linguistic capabilities of AI tools such as LLMs can extend support for struggling readers and writers who require help in decoding, comprehending, and communicating in the conventions and syntax of a non-native language (Marzuki et al. 2023).

Helping struggling readers and writers is often a major focus of education systems everywhere. Many schools invest heavily in specialized literacy intervention programs, additional textbooks, outside specialists, and remedial software. Reading and writing assistance is often highly labor intensive. Though far more research is needed, AI tools may potentially reduce the intensive one-on-one human support traditionally required for struggling readers and writers, particularly L2 learners or those with disabilities, and might serve as an integral element of structured literacy interventions. From a human resource and financial perspective, these

developments could alleviate financial and human resource shortages in education systems struggling with both.

Benefit 4: AI can tailor learning to each student's needs

Education is often maligned for its “one size fits all approach” to teaching and learning. For study participants, AI is changing that paradigm through two primary mechanisms: personalized learning programs and ITSs, which they rank as an important benefit of AI.

PERSONALIZED LEARNING PROGRAMS

Personalized learning tailors instruction to each student's strengths, needs, and interests while providing choice, voice, and flexibility in reaching learning outcomes (Slocum 2016). Although personalization doesn't require technology, the logistical burdens of addressing every student's unique needs means that teachers increasingly *employ* technology, such as AI-powered platforms using LLMs or chatbots to adapt curricula, pacing, and learning environments to individual student profiles.

Teachers in the U.S., Nigeria, India, Singapore, and South Korea interviewed for this study report leveraging AI primarily for strategic adaptations of materials, instruction, and assessments; the tools provide immediate, responsive feedback. These practitioner observations align with research demonstrating that AI technologies facilitate adaptive pacing by responding to individual trajectories, enable continuous assessment with real-time adjustments, enhance engagement through dynamic experiences, and expand opportunities for student questioning (Niemi et al. 2023; Ferman et al. 2021).

Survey data from Cambium Learning Group (2024) confirms this trend: 56% of U.S. teachers currently leverage AI for personalized experiences, 52% employ these tools for real-time performance tracking and feedback, and 50% use AI-powered platforms specifically for developing critical thinking skills, suggesting that personalization extends beyond basic remediation to encompass higher-order cognitive development.

As discussed in Benefit 1, personalized learning can meet diverse learning needs of large student populations with constrained educational resources while offering high-dosage tutoring at low costs (Burns 2021, 37). Personalized systems do this by providing just-in-time feedback, recommendations, explanations, and examples, thereby improving students' conceptual development (Ng et al. 2024, 1333).

Major, Francis, and Tsapali's 2021 meta-analysis of 16 randomized controlled trials across five low- to middle-income countries involving 53,029 learners aged 6–15 years found that technology-supported personalized learning had a statistically significant positive effect size of 0.18 on learning outcomes ($p = 0.001$).

Critically, their meta-regression revealed that adaptive approaches, which adjust to learners' levels, produced significantly greater impact (an effect size of 0.35, considered moderate to strong) than those only linking to learners' interests or providing personalized feedback and assessment (Major et al. 2021, 1935). This demonstrates that adaptive personalization—where instructional content, pacing, or difficulty dynamically adjusts based on current knowledge level or ability—produces substantially stronger learning outcomes than other forms of personalization.

INTELLIGENT TUTORING SYSTEMS: EVOLUTION AND EVIDENCE

In the 1980s, Benjamin Bloom identified education's "2 sigma problem"—namely, that students receiving one-to-one tutoring scored two standard deviations

higher than those without tutoring. The "problem" was making such intensive support scalable and economically feasible (Bloom 1984). Over the past several decades, technology, in the form of computer-based and online tutoring, has been harnessed to provide a scalable, cost-effective alternative to in-person tutoring.

ITSs represent the most common AI educational applications, with millions of users worldwide (Burns 2021). These systems have evolved from early rule-based expert systems using predictive algorithms and knowledge tracing to sophisticated platforms incorporating AI for natural dialogue, dynamic problem generation, and affective data analysis (du Boulay 2016; Buckley et al. 2021). ITSs typically employ cognitive models grounded in learning science theory, using techniques like model tracing or constraint-based modeling to diagnose student knowledge states and provide targeted instructional interventions.

Research consistently demonstrates their effectiveness. Kulik and Fletcher's (2016) meta-review of 50 studies concluded ITSs can "match the success" of human tutoring, while Nickow et al. (2020) found an average effect size of 0.37 standard deviations across 96 randomized controlled trials—equivalent to moving students from the 50th to 65th percentile. Benefits prove strongest in foundational subjects and are particularly helpful for traditionally underserved populations, including learners with disabilities (Kasneci et al. 2023).

While recent ITSs have begun incorporating machine learning methods—including neural networks and natural language processing—to improve student dialogue and predict learning trajectories, most deployed systems with established research evidence still depend on structured, rule-based cognitive architectures from the 1980s and 1990s (García-Méndez 2025). The introduction of LLMs marks a significant new phase in computer-based tutoring, expanding possibilities for flexible conversations and content generation. Granted, research supporting large-scale,

classroom-tested implementations of AI tutoring programs remains limited, lagging behind traditional ITSs that have demonstrated effectiveness through decades of rigorous evaluation (Nye et al. 2023, 2; Jiang et al. 2025). However, three recent studies offer early evidence of the considerable promise that AI-enhanced tutoring systems may hold.

In the first, researchers conducted a randomized controlled trial with 194 undergraduate physics students at Harvard University, comparing learning outcomes and student perceptions between an AI tutor and active learning classroom instruction (Kestin et al. 2025). Both instructional approaches were designed using identical pedagogical best practices. Students alternated between conditions across two weekly lessons on surface tension and fluid flow, completing pre- and post-tests to measure content mastery alongside surveys assessing their learning experience.

Students using the AI tutor demonstrated significantly higher learning gains in less time than those in active learning classrooms. The AI group achieved a median post-test score of 4.5 compared to 3.5 for the classroom group—representing more than double the learning gains relative to the combined pre-test baseline of 2.75. Students also reported feeling more engaged and motivated when working with the AI tutor. The authors conclude that students interacting with the AI tutor at home “learn significantly more than when they engage with the same content during an in-class active learning lesson, while spending less time on task,” suggesting substantial potential for AI tutors in authentic educational settings (Kestin et al. 2025, 4).

The second study, a 2024 World Bank trial in Nigeria found that Microsoft Copilot (generative pre-trained transform, version 4 [GPT-4]) improved first-year secondary students’ English skills by 0.23 standard deviations and digital literacy by 0.31 standard deviations—learning gains equivalent to 1.5 to 2 years of regular schooling in just six weeks. Researchers noted that the program succeeded because it complemented the efforts of teachers

rather than replacing them (Di Simone et al. 2025). Finally, Stanford’s Tutor CoPilot, which provides AI-guided support to human tutors, increased student mastery rates by 4 percentage points overall and 9 percentage points for students with lower-rated human tutors (Wang et al. 2025). At a cost of only \$20 per tutor annually, this may represent an important step in solving Bloom’s Sigma 2 problem.

CONCLUSION: THE IMPORTANCE OF “HYBRID VIGOR”

Despite promising outcomes, AI-powered personalization raises substantial concerns among our study participants. The primary apprehension centers on equity: personalization may also function as sophisticated tracking, placing students on limited pathways where basic skills are reinforced to the detriment of more important competencies. AI’s predictive analytics risk creating self-fulfilling prophecies, whereby algorithmic predictions based on early performance or demographic factors shape educational opportunities in ways that confirm biases. Several Delphi panelists expressed worry that the poorest education systems will employ AI to teach students while students in wealthy education systems will be taught by AI—and human teachers.

Design limitations compound these worries. Because of ease of design, many AI-driven tools employ outdated theories of learning, such as behaviorism, thus limiting both what and how students learn (Chen et al. 2025; Bali 2024). While AI can produce vast quantities of materials leveled to a student’s learning ability, it often lacks the nuanced judgment needed for complex tasks like teaching reading comprehension or writing—where, in the words of one teacher, “human insight is irreplaceable.”

The real danger lies in using tutoring and personalization systems for substitution versus supplementation. When systems position learners as isolated individuals interacting with platforms, they *automate* rather than *augment*

learning, eliminating the embodied experiences and relationships so critical to human learning (Cukurova 2024). AI cannot discern whether student slowness indicates deep thinking, peer collaboration, fatigue, or disengagement—a contextual understanding that good educators develop through daily interaction.

Furthermore, the relational and affective dimensions crucial to genuine personalization—formative feedback, motivational scaffolding, shared goal setting—remain beyond algorithmic capability. Students derive motivation from interpersonal connections; if teachers become sidelined in favor of AI tools, in the words of one Delphi panelist, students may feel devalued as “algorithmic data points rather than individuals whom teachers genuinely care about.”

The evidence suggests AI-based personalization and tutoring work most effectively through human-AI collaboration. One teacher explains what this looks like in practice: “The students come back to me in the small group rotation. The personalized platform has helped them with capitalization, punctuation, and grammar. Now I can help them extend their thinking further.”

This “hybrid vigor” ensures that technology augments, rather than replaces, the human dimensions of meaningful learning. At its best, AI-powered personalization helps each student feel their learning is meaningful while providing teachers with feedback on student progress to enable tailored instruction. The most sophisticated AI-powered methods require a strong teacher-student dialogue to succeed, maintaining the relational foundation essential to education while leveraging technology’s capacity for scale and individualization.

Benefit 5: AI can extend learning to neurodivergent students and students with disabilities

Education systems often struggle to meet the unique needs and strengths of all students. When schools are not accessible and inclusive, participation and learning become especially challenging for students with disabilities and neurodivergent learners. Children with disabilities are nearly 50% more likely to have never attended school. Of the 240 million children with disabilities worldwide, more than 90% in low- and middle-income countries lack access to the assistive technology they need (United Nations Sustainable Development Group 2024). Since most schools are designed for neurotypical children, neurodivergent students—such as those with dyslexia, attention deficit hyperactivity disorder, and autism—are often left out (Teach Access 2025).

AI can help education systems better support children with special needs by expanding the capacity and availability of assistive technology. Globally, more than 2.5 billion people, including both adults and children, need assistive devices such as wheelchairs, hearing aids, or apps that support communication and cognition. Yet nearly one billion lack access, particularly in low- and middle-income countries where availability meets as little as 3% of need (World Health Organization and UNICEF 2022). As we examine next, AI offers promising solutions to address this accessibility gap through applications ranging from adaptive assessments to innovative delivery mechanisms.

and learning from mistakes. This concern emerged even among teachers who have fully integrated AI into their teaching.

Lisanne Bainbridge, a pioneer in automation studies, argued in 1983 that automating most work while leaving humans responsible for non-automatable tasks creates severe problems. As mechanization advances, humans lose proficiency in automated skills no longer exercised regularly, shifting from active skill application to passive system monitoring. Counterintuitively, operators then need more training and different skills for rare but crucial interventions (Bainbridge 1983). Based on interviews with students and teachers, this automation paradox appears evident in schools today. AI use may paradoxically dilute the very competencies necessary for its productive use. As one teacher asked rhetorically, “When you never make mistakes, when you never have failures or challenges, how do you build your resilience and perseverance?” Like Bainbridge’s operators, students risk losing practice with fundamental cognitive and durable skills as they increasingly rely on AI to perform learning’s intellectual work.

DECLINING CRITICAL THINKING AND CREATIVITY

Among all study participants, critical thinking emerges as the predominant concern regarding AI’s impact on students’ cognitive development. Gerlich defines critical thinking as the ability to:

Analyze, evaluate, and synthesize information to make reasoned decisions through cognitive processes including problem-solving, decision-making, and reflective thinking.

This fundamental skill proves essential for academic success, professional competence, and informed citizenship, enabling individuals to process information effectively and engage in reflective thought.

— (Gerlich 2025, 2–3).

Research employing multiple data analysis methods consistently indicates that increased AI tool use

correlates with declining critical thinking skills and less inclination to engage in higher-level thinking (Gerlich 2025; Wahn et al. 2023; Lee et al. 2025; Bastani et al. 2024; Kosmyrna et al. 2025; Sparrow et al. 2011). Again, this outcome aligns with cognitive offloading, whereby students rely on AI to do their thinking instead of exercising analytical reasoning abilities. Similar to muscles atrophying without exercise, critical thinking skills weaken from lack of practice when AI consistently handles complex tasks (Risko and Gilbert 2016).

MIT researcher Nataliya Kosmyrna and colleagues divided 54 college students into three groups, monitoring brain activity via electroencephalography (EEG) headsets across four sessions while students wrote 20-minute essays. The first group used no technology, the second used Internet searches, and the third used ChatGPT. Participants relying on ChatGPT performed worst, showing weaker brain connectivity in regions associated with memory and creativity. Their essays received lower-quality ratings, and showed “diminished critical inquiry, increased vulnerability to manipulation, decreased creativity, and the internalization of shallow or biased perspectives” (Kosmyrna et al. 2025). Kosmyrna, as mentioned at the beginning of this section, called this outcome “cognitive debt,” defining it as a condition where users defer short-term mental effort through “repeated reliance on external systems like LLMs” to replace “effortful cognitive processes required for independent thinking” (Kosmyrna et al. 2025, 141).

Far from unleashing creativity, unfettered LLM use appears to sap it. In Kosmyrna’s study, ChatGPT group essays showed less word choice diversity and were judged “soulless.” Human evaluators questioned whether essays from the ChatGPT group came from the same student. Teachers in our study regularly spoke of AI producing a “homogeneity of ideas,” with students submitting AI-generated answers that were “lifeless, generic, homogenous...lacking any diversity of ideas” and signaling a “loss of creativity” or “originality” caused by “students using AI to think for them.”

BOX 7

Reducing the tendency to offload thinking to AI

AI use need not result in cognitive offloading or decline. Many students already use LLMs to enhance rather than replace their learning, and educationally designed tools like Khanmigo employ question prompts and scaffolding to increase learner participation and discourage passive consumption (Khan Academy 2025).

Several AI companies that provide LLMs have introduced specialized “education modes”—including Guided Learning in Gemini, OpenAI’s Study Mode, and Anthropic’s Claude for Education—that function as personalized learning companions. These configurations prioritize probing questions, step-by-step reasoning guidance, level-adapted explanations, and multimedia reinforcement over direct answers (OpenAI 2025d). While students can still circumvent these safeguards by reverting to instant-answer modes, these developments represent meaningful progress toward supporting active learning.

Teachers can use scaffolds and frameworks to help students reflect on and calibrate their cognitive offloading. The AI Assessment Scale, for example, helps students visualize the boundary between appropriate and inappropriate AI use (Perkins et al. 2024). Similarly, the International Baccalaureate’s *Machine Agents in the Learning Scale* (MAiLS) framework illustrates how cognitive offloading varies with tool agency—from manual tools like pencils to heavily automated ones like chatbots (Quezzaire 2025). As tool agency increases, human agency typically decreases, clarifying the learner-AI relationship and its consequences.

Through frameworks like MAiLS, students can see that low-level AI use (such as online searching or spell checking) poses minimal cognitive offloading concerns, since they still require active thinking. In contrast, medium-level applications like project management automation, database manipulation, or data visualization involve greater offloading and may need deliberate balance with human-centered learning approaches (Quezzaire 2025).

A final simple intervention involves reframing students’ relationship with AI itself, reconceptualizing it not as a *servant* that simplifies work but as a *sparring partner* that sharpens ideas, deepens reasoning, and strengthens learning. This shift helps ensure AI functions as a catalyst for learning rather than an impediment to it (Kentz 2025).

Looking at LLM use across large student cohorts, researchers at Georgetown University warn that AI could reduce collective creativity—the shared pool of creative ideas contributed to society. One study analyzed 2,200 U.S. college admissions essays submitted to undergraduate programs and compared them to essays generated by ChatGPT

using the same prompts (Moon et al. 2025a). Both human- and LLM-generated essays were evaluated on a creativity measure that captured the variety and uniqueness of ideas. In theory, the larger the sample, the more diverse the collective set of ideas should be. However, this was far less true for LLM-generated essays. Researchers found that

each additional human essay expanded the number of unique ideas between two and eight times more than each additional LLM-generated essay (see Figure 10). In a follow up study on 160,000 undergraduate admissions essays, researchers found that “AI-assisted essays and essays written after the release of ChatGPT showed greater lexical diversity, but the increasingly diverse words actually expressed increasingly homogeneous ideas both within and across essays,” they concluded (Moon et al. 2025b).

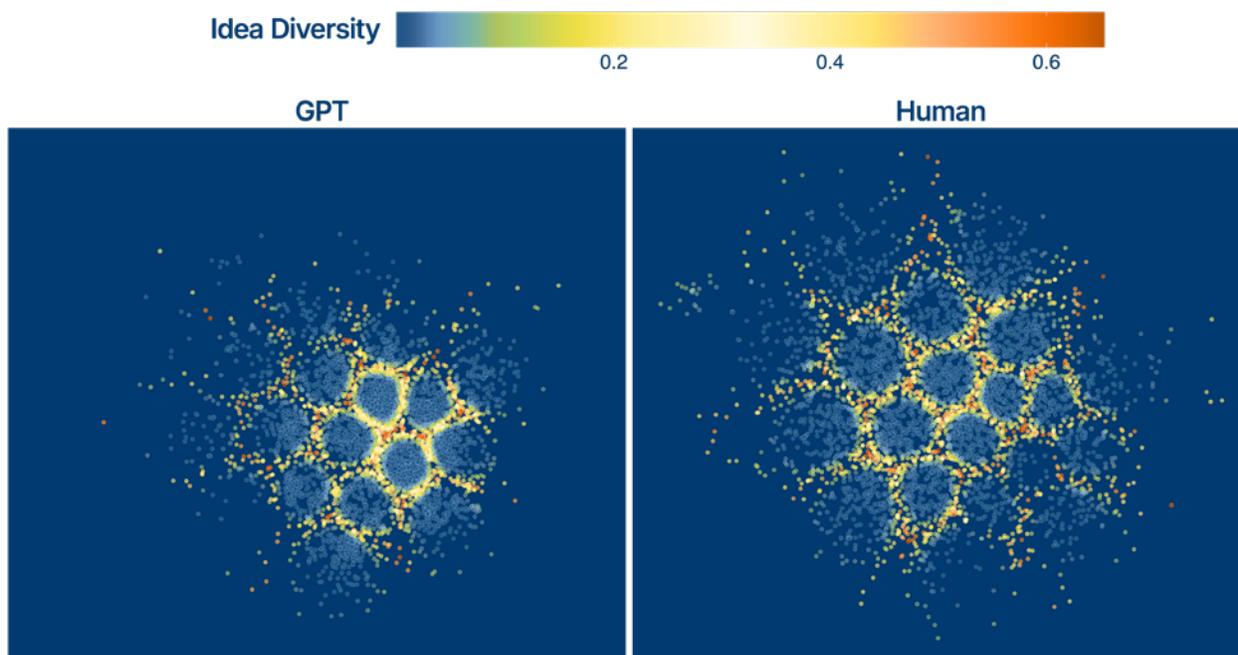
Part of the problem has to do with the design of LLMs themselves and how students approach AI’s outputs. Cognitive science demonstrates that effective learning requires “grounding” information through multiple perspectives, actively integrating new knowledge with existing understanding rather than adopting ready-made responses (Lee et al. 2025). However, AI data relies on “source destruction,” stripping information from larger contexts and creating averaged data that “flattens” the complexity of human knowledge and eliminates

the contextual understanding essential for true learning (Epstein 2025). When students rely on AI-generated content as singular sources, they perceive what it generates as “absolute truth,” in the words of one teacher, missing opportunities for knowledge integration and producing “passive ineffective” responses (Lee et al. 2025, 4).

Evidence from teachers and research shows that students often fail to think critically about AI results, assuming evaluation isn’t needed for “simple” or “secondary” tasks such as generating images or charts. Lee et al. surveyed 319 knowledge workers and identified “awareness barriers” inhibiting critical thinking, including perceiving outputs as indirect, simple, or beyond personal expertise (2025). This tendency intensifies when positive past experiences make students less likely to interrogate outputs, leaving them unaware of how AI changes answers to please users—a sycophancy inherent in LLM design that compromises truth for user satisfaction (Furze 2024).

FIGURE 10

Human essays contribute more unique ideas than ChatGPT-generated essays



Source: Moon et al., 2025a.

well-being and critical thinking development (Miao and Holmes 2023). These concerns intensify when considering how AI systems are deliberately built to satisfy users by mirroring tone, reinforcing views, and simulating empathy, creating illusions of connection that may prove difficult to distinguish from genuine rapport.

Such design features prove particularly problematic for younger students or those seeking emotional support, who may gravitate toward AI precisely because it offers undemanding, frictionless interaction that is always available. While many students will recognize that these conversations are algorithmically generated rather than genuine human interactions, vulnerable populations remain at greatest risk. The danger, in the words of one Delphi panelist, is that these young people will “bypass the demanding human work of negotiation, empathy, patience, and navigating discomfort that authentic relationships require, choosing instead the easier path of AI companionship over genuine human engagement.”

Risk 3: AI can degrade trust in education

This lack of general and specific trust—trust in expertise, in information—creates a nihilism and cynicism where no one trusts anything or anyone.

— Teacher

I don't know who to trust.

— Teacher

I promise you that I wrote this text myself with the help of other humans. I promise you that this is a...product of the human mind. But can you be absolutely sure? A few years ago, you could....Things are different (now)....The text you just read might have been generated by an

alien intelligence of some computer.

— Yuval Noah Harari (2024, 214)

One of the greatest casualties of AI may be the way it has further eroded trust within schools and education systems and damaged public confidence in education to meet the needs of learners and a changing society. Relational trust is the respect and regard for others shown through care, integrity, listening, and treating each other as critical assets, and it is imperative for ensuring education institutions and systems meet the needs of learners and their families, schools, and communities (Bryk and Schneider 2002; Morris and Nora 2024). Longitudinal research across hundreds of U.S. schools emphasizes that relational trust between teachers, families, and school leaders is critical in improving a range of academic and nonacademic outcomes (Bryk and Schneider 2002). This trust helps the wide range of actors across students' lives collaborate and/or have a shared vision of how to help them better learn, grow, and thrive.

In our study, the erosion of relational trust emerged as a significant concern for teachers (16% of all responses), experts (14%), and students (11%). Interview transcripts reveal a widening web of mistrust across the educational ecosystem: teachers increasingly doubt that students produce authentic work, while students think the same about their teachers. Teachers distrust the technology companies developing AI tools, technology companies distrust each other, and these high-level suspicions cascade throughout the educational ecosystem. One of AI's greatest casualties may be the trust that ensures young people have what they need in school to meet their needs and prepare them for the future, which sustains faith in educational institutions themselves.

This research reveals that AI integration has created a complex web of mistrust that extends throughout educational relationships. The teacher-student relationship appears particularly vulnerable to this erosion.

TEACHER-STUDENT RELATIONSHIPS

One of the most important interpersonal relationships in students' lives is teachers, who help young people learn knowledge as well as the various skills needed at the different stages of their lives to prepare them for their futures. Students must believe that teachers care about them, fulfill their professional responsibilities, and possess content expertise and pedagogical competence. Teachers must trust that students' work legitimately represents their abilities and potential.

Trust appears to be dissolving between teachers and students in the age of AI. A 2024 study from the CDT found that more than half of U.S. grade 6–12 public school teachers reported that AI made them “more distrustful” of the integrity of student work (Dwyer and Laird 2024, 12). A 2025 survey by CDT suggests that this erosion correlates with usage patterns—students who use AI for “many” tasks report feeling “more disconnected” from their teachers (56%) compared to students who use it for “few tasks” (46%) (Laird et al. 2025).

AI also compromises teachers' abilities to effectively assess their students' knowledge, skills, and capabilities, which is essential to promoting educational growth and development. AI also undermines teachers' abilities to identify learning difficulties, ensure that students have acquired the foundational skills needed to progress, and design meaningful and tailored instruction, because they don't know what students actually know.

The mistrust flows in both directions. In a U.S. survey, half of students and parents question whether teachers who use AI are “doing their job” (Laird et al. 2025, 21). Both teachers and students perceive AI use by the other party as a “lesser form of care and attention” and disengagement from professional responsibilities (Grose 2025; Burns 2024; Barrett and Pack 2023).

Research indicates that teachers and students find using AI for their respective tasks to be acceptable when how and why they are using it

is transparent and connects to better teaching and learning (Barrett and Pack 2023). Students generally supported teachers using AI for planning but opposed it for creating feedback, and viewed AI-generated grading as signaling that their work is not worth personal engagement. However, a lack of acceptable use policies (AUPs), combined with a perception that AI-generated products are inferior—what Harvard Business School researchers call the “competence penalty”—encourages both groups to hide their AI use (Acar et al. 2025). This teacher's quote captures that tension.

I find that I use AI quite a bit to help make presentations and make worksheets for students. And now I'm kind of thinking...is that fair for me to use these tools to help me in my practice, but maybe say, “Hey, students, you can't use these tools to help you?”

AI disrupts teacher-student relationships along two dimensions. First, students' unauthorized use of AI shifts the relationship from guidance and mentorship toward detective work, as teachers increasingly function as investigators determining work authenticity rather than facilitating learning. Efforts to determine authenticity generate misunderstandings and friction and create a power struggle. Students report frustration at false accusations (Gorichanaz 2023), and one principal interviewed notes that plagiarism detection tools regularly produce false positives. Without formal policies, students can use AI with impunity while teachers feel powerless to respond, further corroding trust and calling into question the legitimacy of grades. This adversarial dynamic risks casting teachers and students as opponents rather than partners in learning (Holzer and Daumiller 2025). Teachers also reported frustration that AI undermined their knowledge and influence, that they are “being pitted against ChatGPT in terms of who has authority.” According to some teachers, students viewed AI as “more authoritative than teachers” and “trust tools but not people”—an observation that, if accurate, fundamentally shifts beliefs about educational authority and credibility.

Second, even when students use AI tools appropriately for seeking help, the relational dynamics shift in consequential ways. Help-seeking behaviors and help-providing behaviors mutually influence each other, shaping trust and rapport between students and teachers. When students redirect their help-seeking from teachers to LLMs, this diminishes the mentor-protégé exchange that is foundational to the teacher-student relationship and central to the professional fulfillment that teachers feel when guiding young learners through intellectual challenges. It also removes a critical opportunity for building relationships and trust between teachers and students, and a touchpoint where teachers and students can recognize each other's assets and contributions to the teaching and learning process.

STUDENT RELATIONSHIPS

AI also appears to be weakening peer interactions and bonds of trust among students (Hou et al. 2025). Instead of interacting with or asking for help from their classmates, students increasingly

rely on AI tools. This failure to interact with other students—especially for assistance, collaboration, or consultation—weakens the reciprocity that is fundamental to student interactions. Over time, if these one-directional exchanges with LLMs persist, the classroom risks losing the spirit of mutual intellectual exchanges and obligations for support that constitute a learning community.

In and of itself, this dynamic can harm trust. Elise Silva and her team (2025) at the University of Pittsburgh conducted focus groups with 95 students about their experiences with AI in academic settings. Students in this study described how AI usage can sow distrust and frustration among peers. Silva describes, “Some talked about asking classmates for help, only to find that they ‘just used ChatGPT’ and hadn’t learned the material. Others pointed to group projects, where AI use was described as ‘a giant red flag’ that made them ‘think less’ of their peers.” The prevalence of delegating work to AI also breeds resentment among students who “play by the rules” and don’t use AI toward

BOX 10

Building trust around AI

The web of mistrust around AI can be ameliorated through multiple actions:

Education systems, with significant representation by teachers and students, can create and implement AUPs outlining appropriate uses of AI. As AI's capabilities evolve and increase, these AUPs can be revisited and modified accordingly. As will be discussed in the Section VI, AI literacy classes can support these AUPs by cultivating academically appropriate and ethical uses of AI by teachers and students.

In addition to AUPs, teachers can establish classroom AI usage guidelines in their classrooms. One model is an AI “traffic light” approach for assignments and projects: green encourages students to use AI; yellow requires teacher preapproval; and red disallows AI use (Mormando 2023).

Teachers can model transparent use of AI showing students how they use it for their own instructional planning, teaching, and assessment, and solicit student input on such use. Teachers and students can examine LLM outputs, critiquing these as stand-alone artifacts and in comparison, with human-generated work. Teachers can require students to document their AI use by including prompts, AI-generated content, verification of information, critical analysis of outputs, and evidence of revision as part of assignments.

those who use AI for their classwork, especially when the latter group receives better grades.

Student interviews highlighted some students' frustration with unclear policies about AI usage, worried that their classmates would simply use chatbots to get better grades than them, and anxious about being falsely accused of AI use. In another study, students were worried that their project teammates would use AI and get their whole project flagged as AI-generated (Gorichanaz 2023). Like much in the world of AI in education, more research is needed on how AI mediates and impacts peer interactions and the social nature of learning (Hou et al. 2025).

TRUST OF AI OVER HUMANS

How AI impacts trust and relationships between teachers and students, students and peers, and teachers and parents/caregivers in educational settings reveals a number of perplexing contradictions. Most study participants expressed deep distrust about the quality of AI tools but were enthusiastic users nonetheless. Students interviewed for this study expressed deep distrust of AI systems due to LLMs' tendency to "hallucinate"—a concern substantiated by OpenAI's study on its own reasoning models, which documented hallucination rates ranging from 33% to 79%, depending on the test (Weissberger 2025). Yet, paradoxically students appear to use AI outputs with limited oversight, despite their stated skepticism (Mah et al. 2025). This contradiction becomes particularly concerning as AI amplifies the spread of misinformation through hallucinations, which students tend to consume uncritically (Fukuyama 2025), while many remain uncertain whether they are interacting with humans or chatbots in digital environments.

This dichotomy may stem from two sources. First, discussed in Risk 2, deception is intrinsic to the design of technology tools. Students may interact with AI "believing they're engaging with something approaching human-level intelligence, when they're actually using a sophisticated pattern-matching

system" (Potkalitsky 2024). The human tendency to anthropomorphize combined with LLM sycophancy, technical complexity, and lack of AI literacy courses may discourage any detailed scrutiny of AI responses by students. This results in greater trust in AI versus humans, less oversight of AI outputs, greater student disclosure to AI avatars, and increased vulnerability to manipulation and reliance on AI tools (Ghosh 2025; Lee et al. 2025).

As noted earlier in Risk 1, teachers consistently observe that students trust AI outputs uncritically. The students in this study, who are in secondary school, are still relatively young and lack the deep content knowledge necessary to ascertain the veracity of AI-generated information. Teachers expressed particular concern that as AI-generated content becomes more widespread, students' ability to distinguish fact from fiction will become even more compromised.

Second, both students and teachers may be experiencing fundamental erosion of self-trust—diminishing confidence in their own intellectual capabilities as they increasingly delegate cognitive work to AI. Research suggests that students fail to disclose AI use for fear of being judged as "lazy," "stupid," or "foolish," and skepticism toward AI users is common (Hou et al. 2025, 4) This loss of faith in human judgment and ability may represent the most profound trust crisis of all, as it undermines the very foundation of learning and personal development that education seeks to foster (Hou et al. 2025, 6).

DISTRUST OF "BIG TECH"

Despite their enthusiasm about AI, study participants express the highest degree of distrust toward "big tech" companies, in particular AI technology developers creating consumer products. Almost all participants expressed significant distrust of these companies, their leaders, their motivations, their interactions with schools, and the AI tools themselves—even while using them. They are not alone. Though it focuses on social media companies, a 2024 Pew survey