AI Tutors in Education: Design Assumptions, Learning Theory, and Instructional Impact

Category: AI and Education | Published: February 16, 2026

Dylan Bulseco

Institute for Future Intelligence, Natick, MA, USA. dylan@intofuture.org

Abstract

AI tutoring systems are increasingly adopted to address long-standing educational challenges related to scale, individualization, and instructional consistency. This blog examines AI tutors through the lens of learning science and instructional design, rather than treating advances in artificial intelligence as inherently educationally beneficial. We trace the evolution of AI tutors from early Intelligent Tutoring Systems to contemporary generative, dialogue-based models, emphasizing how changes in system architecture alter both instructional affordances and pedagogical risks. Focusing on mechanisms of personalization, we analyze how adaptive sequencing, feedback and scaffolding, pacing, and content customization are implemented in current AI tutoring systems, and how these mechanisms align with established theories of learning. Drawing on empirical and conceptual literature, we evaluate evidence related to learning outcomes, engagement, self-regulated learning, and equity. We argue that reported benefits are strongly contingent on domain structure, instructional design, alignment between tutoring activities and assessment, and learners' prior knowledge and self-regulation skills. In particular, gains are most consistently observed in structured domains and short-term outcomes, while evidence for transfer and durable conceptual understanding remains limited. This blog further treats ethical and institutional considerations, including bias, data governance, and changes to teacher roles, as integral design constraints rather than external concerns. We conclude that AI tutors should be understood as instructional tools embedded

within sociotechnical systems. Their educational value depends on theoretically grounded design, careful classroom integration, and evaluation practices that prioritize meaningful learning over surface-level performance.

Introduction

The field of education is undergoing a profound transformation, driven by the rapid advancements and increasing integration of Artificial Intelligence (AI) (Wang, 2024; Chen, 2020). While AI has been a subject of research in education for decades, the recent high performance of generative AI models like ChatGPT on standardized academic tests has propelled the topic into the mainstream, sparking widespread discussion about the future of teaching and learning (Kamalov, 2023). This has led to an exponential growth in research and development focused on AI in educational settings, marking the dawn of a new era for educational technology (Bahroun, 2023).

Traditional educational models have long been characterized by a "one-size-fits-all" approach, where a single teacher delivers a standardized curriculum to a diverse group of students. This method often struggles to accommodate the wide variations in prior knowledge and learning styles of individual students (Pratama, 2023). Such systems, often teacher-centered and focused on memorization, can fail to adequately engage all learners or cater to individual needs, potentially widening achievement gaps (Banik, 2025). The ideal of personalized education, where the learning experience is tailored to each student, has historically been a resource-intensive goal, achievable only through one-on-one human tutoring.

AI offers a scalable solution to this challenge. By leveraging technologies such as machine learning and natural language processing, AI can create adaptive and personalized learning environments (Harry, 2023). Among the most promising applications of AI in education (AIEd) are AI-powered tutoring systems, often referred to as Intelligent Tutoring Systems (ITS) or AI Tutors. These systems are designed to emulate the benefits of a human tutor, providing individualized instruction, guidance, and feedback to students as they navigate learning materials.

This blog aims to provide a comprehensive review of the rise of AI tutors in modern education. We will explore their technological evolution, the mechanisms they employ to achieve personalization, and their demonstrated impact on student learning and engagement. Furthermore, we will critically examine the significant challenges, ethical considerations, and implementation hurdles that must be addressed for their successful integration. By

synthesizing the existing literature, this paper seeks to provide a nuanced understanding of how AI tutors are reshaping the educational landscape and to identify key directions for future research and practice to ensure their responsible and effective deployment.

The Evolution of AI Tutors: From Intelligent Tutoring Systems to Generative AI

The aspiration to use technology for individualized instruction long predates contemporary artificial intelligence. Early efforts in computer-assisted instruction (CAI) during the mid-20th century focused primarily on automating drill-and-practice exercises and linear content delivery. While these systems improved access and consistency, they lacked the capacity to reason about learner understanding or adapt meaningfully to individual needs. The limitations of CAI motivated a shift toward more cognitively informed approaches, laying the foundation for what would become modern AI tutoring systems.

Intelligent Tutoring Systems: Structured Adaptivity

The first true precursors to today's AI tutors were **Intelligent Tutoring Systems (ITS)**, which emerged from artificial intelligence and cognitive science research in the 1970s and 1980s (VanLehn, 2011). Unlike CAI, ITS were explicitly designed to model not only instructional content but also learner cognition and instructional decision-making. Seminal work in this area drew heavily on theories of human problem solving and expertise, particularly in well-structured domains such as mathematics, physics, and computer programming (Anderson et al., 1995; Woolf, 2010).

A canonical ITS architecture consists of three interdependent components:

1. **The domain model**, which encodes the knowledge, skills, and problem-solving strategies to be taught;
2. **The student model**, which represents the learner's current knowledge state, misconceptions, and progress over time; and
3. **The pedagogical (or tutoring) model**, which governs instructional decisions such as task selection, feedback timing, and remediation strategies (Göksel, 2019; Alcívar, 2024).

This architecture enabled ITS to deliver adaptive instruction that approximated some benefits of one-on-one human tutoring, an effect famously articulated in Bloom's "2 sigma problem," which highlighted the dramatic learning gains achievable through individualized tutoring compared to conventional classroom instruction (Bloom, 1984). Decades of empirical research demonstrated that well-designed ITS could significantly improve learning outcomes, particularly for procedural and rule-based knowledge (VanLehn, 2011). However, their strengths were also their constraints: ITS required extensive domain engineering, were costly to develop, and struggled to generalize beyond narrowly defined problem spaces or predefined interaction patterns.

The Shift to Generative AI and Large Language Models

The most recent—and arguably most disruptive—phase in the evolution of AI tutors has been driven by advances in **Generative AI (GenAI)**, particularly large language models (LLMs). The public release and rapid adoption of systems such as ChatGPT marked a qualitative shift in what AI tutors could plausibly do in educational contexts (Bahroun, 2023). Rather than relying primarily on hand-crafted rules or expert-authored knowledge representations, LLMs leverage large-scale pretraining on natural language corpora to generate context-sensitive responses across a wide range of topics.

This represents a fundamental departure from classical ITS. GenAI-powered tutors are capable of engaging in open-ended, conversational interactions; generating multiple explanations tailored to a learner's framing of a question; producing novel examples or practice problems on demand; and interpreting ambiguous or incomplete student inputs with far greater flexibility than earlier systems (Leon, 2024). As a result, the interaction paradigm has shifted from system-directed tutoring toward **learner-initiated, dialogue-driven learning**, lowering barriers to entry and broadening applicability across disciplines.

From the learner's perspective, these systems are no longer experienced merely as instructional tools but as **multi-functional learning partners**. Students increasingly use GenAI systems as writing collaborators, conceptual explainers, exam-preparation coaches, and even simulated peers for brainstorming and rehearsal (Kim, 2024). Their persistent availability enables on-demand academic support unconstrained by classroom schedules or institutional resources, effectively extending learning opportunities beyond traditional educational settings (Arya, 2024; Jian, 2023).

Retrieval-Augmented Generation: Addressing the Limits of Pure Generation

Despite their power, purely generative LLM-based tutors suffer from well-documented limitations, most notably hallucinations, opaque reasoning processes, and reliance on static, parametric knowledge that may be outdated or misaligned with a specific curriculum. **Retrieval-Augmented Generation (RAG)** has emerged as a critical architectural advance to address these shortcomings, particularly for knowledge-intensive tasks such as tutoring and assessment (Lewis et al., 2020).

RAG systems combine generative models with an external retrieval mechanism that dynamically accesses curated knowledge sources at inference time. This approach reduces factual errors, improves grounding in authoritative materials, and allows instructional content to be updated or localized without retraining the underlying model. For education, this has several important implications: it enables alignment with specific textbooks or curricula, supports domain specialization through vetted knowledge bases, and enhances transparency by making it possible to inspect the sources informing a tutor's responses.

In principle, RAG also overcomes a fundamental scalability constraint of traditional ITS and standalone LLMs by decoupling reasoning and language generation from long-term knowledge storage. This makes it particularly well suited to evolving curricula, interdisciplinary learning, and rapidly changing fields where static knowledge representations quickly become obsolete.

However, these advantages are contingent on rigorous system design. Poor-quality embeddings, weak retrieval strategies, or inadequate re-ranking mechanisms can introduce new failure modes, including selective bias, incomplete evidence retrieval, or misplaced confidence in incorrect sources. Moreover, without careful prompting and evaluation, RAG systems may still generate responses that diverge from retrieved evidence, undermining their pedagogical reliability. Consequently, the effective deployment of RAG-based AI tutors requires systematic evaluation, iterative refinement, and close collaboration between educators, learning scientists, and AI engineers.

From Engineered Tutors to Adaptive Learning Ecosystems

Taken together, the evolution from CAI to ITS and now to GenAI-powered, RAG-enhanced systems reflects a broader shift in educational technology: from tightly engineered instructional artifacts toward flexible, data-driven learning ecosystems. While generative AI dramatically expands the scope and accessibility of personalized learning, it also reintroduces

enduring questions about accuracy, trust, and pedagogical intent. Understanding this historical trajectory is essential for evaluating not only what AI tutors can do, but how, and under what conditions, they should be integrated into education at scale.

Mechanisms of Personalization in AI Tutoring Systems

Personalization in AI tutoring systems is often described as an inherent or automatic property of artificial intelligence, but in practice it is the result of a series of concrete design decisions about what data to collect, how to interpret learner behavior, and which instructional actions to prioritize. Not all AI tutors are equally personalized, and not all forms of personalization are pedagogically beneficial. Understanding how personalization is operationalized is therefore essential for evaluating both the promise and the limitations of these systems.

Adaptive Learning Paths: Strengths and Constraints

One of the most visible personalization mechanisms in AI tutors is adaptive sequencing of content. In principle, adaptive learning paths allow students to bypass material they have already mastered while receiving additional support on concepts they find difficult. When implemented carefully, this approach can approximate mastery-based learning by aligning progression with demonstrated understanding rather than fixed timelines.

In practice, however, adaptivity is only as good as the model of learning that underlies it. Many systems infer mastery primarily from short-term performance indicators such as correctness or response time. These proxies can be misleading. For example, a student may answer correctly through pattern recognition or guesswork without developing a robust conceptual understanding. Conversely, productive struggle—an important driver of learning—may be misclassified as failure, triggering unnecessary remediation. As a result, adaptive paths can sometimes optimize for efficiency rather than learning depth, accelerating students through content before their understanding is stable.

Feedback and Scaffolding: When “Immediate” Is Not Always Better

AI tutors are particularly effective at delivering rapid, targeted feedback. Unlike classroom settings where feedback may be delayed or unevenly distributed, AI systems can respond to each learner action in real time. This capability is especially valuable for procedural domains, where stepwise feedback can prevent the accumulation of errors.

That said, immediacy is not inherently beneficial. Educational research has long shown that overly frequent or overly explicit feedback can reduce opportunities for reflection and self-correction. Some AI tutors inadvertently short-circuit learning by intervening too quickly, offering hints or corrections before students have had time to reason through a problem. The challenge, therefore, is not simply to provide feedback, but to calibrate its timing, specificity, and instructional intent. Human tutors routinely make these judgments implicitly; encoding them into AI systems remains an open pedagogical problem rather than a solved technical one.

Generative Models and the Illusion of Personal Understanding

Recent generative AI–based tutors introduce a qualitatively different form of personalization: conversational adaptation. Large language models can rephrase explanations, generate examples on demand, and respond flexibly to student questions. From the learner's perspective, this can feel like being understood at a personal level.

However, this perceived personalization is often linguistic rather than cognitive (Chi, 2009). Generative models do not inherently maintain a stable or theory-driven representation of a learner's conceptual state. Without explicit student modeling or external constraints, the system may generate fluent explanations that are misaligned with what the learner actually needs. In this sense, conversational personalization can create an illusion of individualized instruction while masking shallow diagnostic depth. Hybrid approaches that combine generative interaction with explicit learner models or retrieval-based grounding are therefore more promising than generation alone.

Pacing, Autonomy, and the Risk of Self-Regulation Gaps

Allowing learners to control pacing is frequently cited as a major advantage of AI tutors. Self-paced instruction can reduce anxiety, accommodate diverse schedules, and empower students to take ownership of their learning. For motivated and metacognitively skilled learners, this flexibility can be transformative.

Yet pacing control also shifts responsibility from the instructor to the learner, and not all students are equally prepared for this shift. Learners with weaker self-regulation skills may procrastinate, disengage, or repeatedly avoid challenging material. AI tutors can partially mitigate this risk by monitoring engagement patterns and prompting reflection or goal-

setting, but such interventions are still rudimentary compared to human judgment. Personalization that increases autonomy without supporting self-regulation may inadvertently widen achievement gaps rather than close them.

Content Customization: Engagement Versus Coherence

AI tutors can tailor examples, analogies, and contexts to student interests, which often increases engagement and perceived relevance. For instance, framing math problems around sports, games, or career interests can lower affective barriers to learning.

However, excessive customization carries its own risks. When instructional content becomes too closely aligned with personal interests, learners may struggle to transfer knowledge to less familiar contexts. There is also a tension between personalization and curricular coherence: highly individualized content can make it difficult to ensure that all learners encounter core concepts, representations, and disciplinary norms. Effective personalization therefore requires careful constraint—deciding not only what to adapt, but what must remain consistent across learners.

Personalization as a Pedagogical Choice, Not a Technical Default

Taken together, these mechanisms illustrate that personalization in AI tutoring systems is not a binary feature but a spectrum of design trade-offs. Decisions about adaptivity, feedback, pacing, and content reflect implicit theories of learning and values about efficiency, autonomy, and equity. Treating personalization as an automatic benefit of AI obscures these choices and risks overestimating instructional impact. A more productive framing views AI tutors as tools that can support personalization under specific conditions, rather than as inherently personalized educators.

Impact on Student Learning and Engagement

Claims about the educational impact of AI tutors often emphasize improvements in learning outcomes, engagement, and skill development. While there is credible evidence that AI-supported instruction can benefit students under certain conditions, the magnitude and durability of these effects vary widely across domains, learner populations, and instructional designs. Understanding when AI tutors help, when they merely assist, and when they may hinder learning requires moving beyond aggregate performance gains to examine underlying learning processes.

Learning Outcomes: Gains, Limits, and Context Dependence

Empirical studies consistently show that AI tutors can improve short-term learning outcomes, particularly in structured domains such as mathematics, programming, and procedural science tasks. These gains are most robust when AI systems provide targeted practice, timely feedback, and adaptive sequencing aligned with clearly defined learning objectives. In such contexts, AI tutors can approximate some of the benefits traditionally associated with one-on-one tutoring, especially for students who would otherwise receive limited instructional support.

However, learning gains reported in the literature are often measured using assessments closely aligned with the tutor's instructional content. This raises questions about transfer and generalization. Improvements in problem-solving performance within the system do not always translate to improved performance on novel tasks or conceptual assessments. In some cases, students learn how to succeed within the tutor rather than developing a deeper understanding of the underlying subject matter. As a result, AI tutors appear most effective as supplements to instruction rather than as standalone learning environments.

Engagement and Motivation: Productive Support or Superficial Stimulation?

AI tutors frequently increase behavioral engagement by offering interactive interfaces, immediate responses, and adaptive challenges. For many students, particularly those who are hesitant to ask questions in classroom settings, the nonjudgmental nature of AI-based support lowers barriers to participation. This can be especially beneficial for learners who struggle with confidence or anxiety in academic contexts.

Yet increased engagement does not automatically imply improved learning. Some AI systems rely heavily on gamification or continuous feedback, which can shift students' focus toward task completion rather than sense-making. When engagement is driven primarily by external rewards or novelty, its effects may diminish over time. Moreover, constant availability of assistance can reduce opportunities for sustained effort and productive struggle, both of which are important for developing expertise. The challenge lies in designing AI tutors that support persistence and curiosity without fostering dependence or passivity.

Higher-Order Skills and the Risk of Cognitive Offloading

A frequently cited advantage of AI tutors is their potential to support the development of higher-order skills such as critical thinking, problem decomposition, and self-regulated learning. In theory, adaptive feedback and reflective prompts can encourage learners to monitor their understanding, revise strategies, and learn from errors.

In practice, however, generative AI tutors introduce a tension between support and substitution. When systems provide complete explanations, worked solutions, or polished text on demand, students may bypass essential cognitive processes. This phenomenon, sometimes described as cognitive offloading, risks weakening learners' capacity to reason independently. Without explicit design constraints—such as requiring justification, prediction, or reflection—AI tutors may inadvertently reward answer acquisition rather than knowledge construction. Evidence suggests that learning benefits are strongest when AI systems are deliberately designed to slow students down at key moments and prompt reasoning rather than efficiency.

Equity, Accessibility, and Differential Impact

One of the most compelling arguments for AI tutors is their potential to expand access to individualized educational support. Students who lack access to human tutors due to cost, geography, or institutional limitations can benefit from on-demand assistance. AI tutors can also support accessibility through multimodal explanations, adjustable pacing, and language support.

At the same time, the impact of AI tutors is not uniform across learners. Students with stronger prior knowledge and self-regulation skills tend to benefit more from open-ended AI support, while those with weaker foundations may struggle to use these tools effectively. Additionally, disparities in access to reliable technology, quiet study environments, and digital literacy can shape who benefits most. Without careful implementation, AI tutors risk amplifying existing educational inequalities rather than mitigating them.

Learner Perception and the Changing Meaning of "Help"

AI tutors also influence how students conceptualize learning and help-seeking. For some learners, AI-mediated assistance normalizes asking questions and experimenting with ideas, fostering a sense of agency. For others, the constant presence of an "answerable" system can redefine learning as a process of querying rather than thinking. Over time, this may alter students' expectations of instruction, feedback, and effort.

These shifts are not inherently negative, but they warrant attention. Educational systems must consider not only whether AI tutors improve performance, but how they shape learners' identities, habits, and relationships to knowledge.

Interpreting Impact with Caution

Taken together, the evidence suggests that AI tutors can positively influence learning and engagement, but only under specific pedagogical conditions. Their effectiveness depends less on the sophistication of the underlying technology than on how well their instructional behaviors align with established principles of learning. Treating AI tutors as universally beneficial obscures important trade-offs and risks. A more defensible conclusion is that AI tutors are powerful educational instruments whose impact is contingent, uneven, and highly sensitive to design choices and contextual factors.

Discussion: Challenges and Ethical Considerations

Despite the immense potential of AI tutors, their widespread adoption is accompanied by a host of significant challenges and ethical considerations that must be carefully navigated by educators, developers, and policymakers. The path to successful integration requires a critical examination of the potential pitfalls.

Algorithmic Bias: One of the most pressing concerns is the risk of algorithmic bias. AI systems learn from data, and if the training data reflects existing societal biases related to race, gender, socioeconomic status, or language, the AI tutor may perpetuate or even amplify these inequities (Bahroun, 2023). A biased system might offer less effective support to certain student populations or use culturally insensitive examples, thereby introducing barriers to already marginalized groups and undermining the goal of equity (Drouet, 2024; Adel, 2024).

Data Privacy and Security: AI tutors function by collecting and analyzing vast quantities of student data, including performance metrics, interaction logs, and potentially even biometric or affective data. This raises serious concerns about data privacy and security (Harry, 2023). Robust policies and technical safeguards are needed to protect this sensitive information from misuse, commercial exploitation, or security breaches. Ensuring transparency in how data is collected, used, and stored is paramount to building trust among students, parents, and educators (Yambal, 2025; Bit, 2024).

Over-reliance and Deskilling: There is a valid concern that students might become overly dependent on AI tutors, using them as a crutch to obtain answers rather than as a tool to facilitate learning. This over-reliance could potentially hinder the development of essential skills such as critical thinking, resilience in problem-solving, and information literacy (Tam, 2023). The design of AI tutors must therefore prioritize pedagogical strategies that encourage deep thinking and productive struggle, rather than simply optimizing for task completion. Equally important is the design of assessment, which must be aligned with these pedagogical goals by rewarding reasoning processes, reflection, and independent synthesis, rather than the mere production of correct or AI-assisted answers.

The Evolving Role of the Human Teacher: The rise of AI tutors inevitably raises questions about the future role of human teachers. While fears of replacement are common, the consensus in the literature points toward a collaborative future where AI augments, rather than supplants, human instruction (Bello, 2025). Indeed, students themselves often do not believe AI can fully replace human educators (Pratama, 2023). In this new paradigm, the teacher's role shifts from being a "sage on the stage" to a "guide on the side." AI tutors can handle the more routine aspects of instruction and practice, freeing up teachers to focus on higher-value tasks such as mentoring, facilitating collaborative projects, fostering social-emotional learning, and providing nuanced, holistic support that AI cannot replicate (Jiménez, 2024).

Implementation Challenges: On a practical level, the widespread and equitable implementation of AI tutors faces significant hurdles. These include the high cost of development and procurement, the need for robust technological infrastructure in schools, and the persistent issue of the digital divide, which could exacerbate inequalities if not addressed (Canonigo, 2024). Crucially, effective integration requires comprehensive professional development to equip teachers with the skills and pedagogical knowledge to use these new tools effectively (Merino-Campos, 2025). Without proper training and support, even the most advanced AI tutors may be underutilized or used in ways that are not pedagogically sound (Sihag, 2024).

Future Directions and the Evolving Classroom

Discussions of the future of AI tutors often emphasize rapid technological advancement, yet meaningful progress in educational contexts will depend less on model capability than on alignment with well-established theories of learning, instruction, and human development.

Future research and implementation efforts should therefore be guided by explicit pedagogical frameworks rather than by assumptions that increased automation or generative flexibility will naturally produce better learning outcomes.

From Standalone Tools to Instructionally Embedded Systems

One of the most important directions for AI tutoring systems lies in their integration into broader instructional ecosystems. Rather than functioning as autonomous learning environments, AI tutors are more plausibly conceptualized as components within coordinated instructional designs that include teachers, peers, curricula, and assessments. From a learning sciences perspective, this aligns with sociocultural theories of learning, which emphasize that knowledge construction is mediated through interaction, norms, and shared practices rather than isolated individual activity.

Embedding AI tutors within classroom routines enables their outputs—such as learner progress data, common misconceptions, and interaction histories—to inform instructional decision-making. However, this integration must be theoretically grounded. Dashboards and analytics should be designed to support teachers' pedagogical reasoning rather than overwhelm them with decontextualized metrics. Research on teacher cognition suggests that actionable insights must be interpretable, timely, and aligned with instructional goals, a standard that many current AI-driven analytics tools do not yet meet.

Grounding Personalization in Models of Learning

Future advances in AI tutoring should move beyond surface-level personalization toward theoretically grounded learner modeling. Many current systems adapt primarily on the basis of performance indicators, such as correctness or completion time, which provide limited insight into conceptual understanding. Cognitive theories of learning emphasize the importance of mental models, misconceptions, and prior knowledge structures—constructs that are difficult but not impossible to approximate computationally.

Progress in this area will require interdisciplinary collaboration between AI developers and learning scientists to operationalize cognitive constructs in ways that are both measurable and instructionally meaningful. Rather than treating personalization as an end in itself, future systems should make explicit claims about which aspects of learning they are adapting to and why. This shift would allow personalization strategies to be evaluated not only in terms of efficiency or engagement, but also in terms of their alignment with how learning is known to occur.

Designing for Metacognition and Self-Regulated Learning

A theory-driven future for AI tutors must also address metacognition and self-regulation, which are central determinants of long-term learning success. While AI tutors can support practice and feedback, their potential contribution to helping learners plan, monitor, and evaluate their own learning remains underdeveloped. Research in self-regulated learning suggests that students benefit from prompts that encourage goal-setting, reflection, and strategic decision-making, particularly when these prompts are embedded within authentic learning tasks.

Rather than passively responding to student queries, future AI tutors could be designed to occasionally withhold assistance, prompt learners to articulate reasoning, or ask learners to predict outcomes before providing feedback. Such design choices reflect theoretical commitments about learning as an active, effortful process. Importantly, these interventions should be empirically evaluated, as poorly timed or overly intrusive metacognitive prompts can be counterproductive.

Reframing the Role of the Teacher Through Theory, Not Replacement

The evolving role of teachers in AI-supported classrooms should be framed through instructional theory rather than narratives of automation or replacement. Research consistently shows that effective teaching involves diagnosing student thinking, orchestrating social interaction, and responding flexibly to classroom dynamics—capabilities that extend beyond current AI systems. AI tutors may reduce the burden of repetitive practice or provide additional support, but they do not eliminate the need for pedagogical judgment.

A more productive research direction examines how AI tutors can redistribute instructional labor in ways that enhance, rather than diminish, teacher agency. For example, when AI tutors handle routine feedback, teachers may have more capacity to engage in formative assessment, small-group instruction, or mentorship. These hypotheses should be tested through classroom-based studies that examine not only student outcomes but also changes in teacher practice and instructional quality.

Assessment, Accountability, and Alignment

Future deployment of AI tutors must also confront issues of assessment and accountability. If AI systems meaningfully shape how students learn, then assessments must be designed to measure the kinds of understanding and skills these systems are intended to support. Misalignment between AI-supported learning activities and traditional assessments risks distorting instructional priorities and incentivizing superficial engagement.

Theory-driven approaches to assessment, such as evidence-centered design, offer a framework for aligning learning objectives, instructional activities, and evaluation. Applying such frameworks to AI tutoring systems would allow educators and researchers to make more principled claims about their educational value and limitations. Without this alignment, claims about impact will remain difficult to interpret and compare across contexts.

Ethical and Institutional Considerations as Design Constraints

Finally, ethical considerations should be treated not as external constraints imposed after development, but as integral components of system design. Issues such as data privacy, transparency, and fairness intersect directly with theories of trust, agency, and institutional responsibility. Future research should examine how design choices influence learner and teacher trust, and how governance structures shape the use of AI tutors in practice.

Rather than asking what AI tutors might become, a theory-driven approach asks what educational purposes they are meant to serve and under what conditions they should be used. This reframing shifts the focus from speculative futures to empirically grounded design questions that can be systematically investigated.

Conclusion

The increasing presence of AI tutoring systems in educational settings reflects a broader shift toward data-informed and adaptive instructional technologies. While these systems offer new capabilities for individualized practice, feedback, and content access, their educational significance cannot be assessed solely on the basis of technical sophistication or availability. The central question is not whether AI tutors can personalize learning, but under what pedagogical assumptions, design constraints, and institutional conditions such personalization supports meaningful learning (Mayer, 2004).

Across decades of research, from early Intelligent Tutoring Systems to contemporary generative models, evidence suggests that AI tutors are most effective when their instructional behaviors are grounded in established theories of learning. Adaptive sequencing, feedback, and scaffolding produce reliable benefits primarily in structured domains and when aligned with clear learning objectives. Conversely, systems that emphasize efficiency, fluency, or answer generation without modeling learner understanding risk promoting shallow engagement and limited transfer. These outcomes highlight a persistent lesson from the learning sciences: instructional effectiveness depends on design decisions informed by theory, not on automation alone.

The introduction of generative AI expands the expressive and interactive capacity of tutoring systems but also introduces new instructional risks. Conversational flexibility and on-demand explanations can enhance accessibility and engagement, yet they can also obscure diagnostic limitations and encourage cognitive offloading if not carefully constrained. Hybrid architectures that combine generative interaction with explicit learner models and grounded instructional content offer a more defensible path forward, though their effectiveness remains an empirical question rather than an established result.

Importantly, AI tutors should be understood as components of sociotechnical systems rather than independent educational agents. Their impact is shaped by how they are embedded within classrooms, curricula, assessment practices, and institutional norms. Human teachers continue to play a central role in interpreting learner behavior, orchestrating social learning, and exercising pedagogical judgment. Research and implementation efforts that frame AI tutors as complements to teaching—rather than substitutes for it—are more consistent with both empirical evidence and educational theory.

Future progress in this area will depend on shifting emphasis from speculative potential to theoretically grounded design and evaluation. This includes making explicit the learning theories that guide personalization strategies, aligning AI-supported instruction with principled assessment frameworks, and examining differential effects across learner populations and contexts. Ethical considerations such as equity, transparency, and data governance should be treated as design requirements rather than post hoc concerns, as they directly influence trust and educational legitimacy.

In sum, AI tutors represent neither a panacea nor a threat to education. They are instructional tools whose value emerges from the interaction between technology, pedagogy, and context. Advancing their responsible use requires sustained collaboration between educators, learning scientists, and technologists, as well as a commitment to evaluating AI tutors not by what they promise, but by how they function within real learning environments.

References

Aleven, V., Blankestijn, J., Lawrence, L., Nagashima, T., & Taatgen, N. (2022). A dashboard to support teachers during students' self-paced AI-supported problem-solving practice. *Lecture Notes in Computer Science*, 16–30. https://doi.org/10.1007/978-3-031-16290-9_2

Bahroun, Z., Anane, C., Ahmed, V., & Zacca, A. (2023). Transforming education: A comprehensive review of generative artificial intelligence in educational settings through bibliometric and content analysis. *Sustainability*, 15(17), 12983.https://doi.org/10.3390/su151712983

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6), 4–16.https://doi.org/10.3102/0013189X013006004

Canonigo, A. (2024). Leveraging AI to enhance students' conceptual understanding and confidence in mathematics. *Journal of Computer Assisted Learning*, 40(6), 3215–3229.https://doi.org/10.1111/jcal.13065

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 1, 100002. https://doi.org/10.1016/j.caeai.2020.100002

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1(1), 73–105. https://doi.org/10.1111/j.1756-8765.2008.01005.x

Göksel, N., & Bozkurt, A. (2019). Artificial intelligence in education. In *Advances in Educational Technologies and Instructional Design* (pp. 224–236). IGI Global. https://doi.org/10.4018/978-1-5225-8431-5.ch014

Jaboob, M., Hazaimeh, M., & Al-Ansi, A. (2024). Integration of generative AI techniques and applications in student behavior and cognitive achievement in Arab higher education. *International Journal of Human–Computer Interaction*, 41(1), 353–366. https://doi.org/10.1080/10447318.2023.2300016

Kamalov, F., Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. *Sustainability*, 15(16), 12451. https://doi.org/10.3390/su151612451

Kim, J., Yu, S., Detrick, R., & Li, N. (2024). Exploring students' perspectives on generative AI-assisted academic writing. *Education and Information Technologies*, 30(1), 1265–1300.https://doi.org/10.1007/s10639-024-12878-7

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, 59(1), 14–19. https://doi.org/10.1037/0003-066X.59.1.14

Shum, S. B., & Luckin, R. (2019). Learning analytics and AI: Politics, pedagogy and practices. *British Journal of Educational Technology*, 50(6), 2785–2793. https://doi.org/10.1111/bjet.12880

Tam, W., Huynh, T., Tang, A., Luong, S., Khatri, Y., & Zhou, W. (2023). Nursing education in the age of artificial intelligence-powered chatbots: Are we ready yet? *Nurse Education Today*, 129, 105917. https://doi.org/10.1016/j.nedt.2023.105917

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221. https://doi.org/10.1080/00461520.2011.611369

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, 252, 124167.https://doi.org/10.1016/j.eswa.2024.124167

© 2026, Dylan A. Bulseco, Ph.D., AI in Education Blog. All rights reserved.

This work is funded in part by NSF DUE 2300796, NSF ITE 2347375 and NSF DRL 2342700.