Building responsible AI chatbot platforms in higher education: an evidence-based framework from design to implementation

Julie Schell[1,2,3,4]*, Kasey Ford[3] and Arthur B. Markman[5]

[1]Department of Design, College of Fine Arts, The University of Texas at Austin, Austin, TX, United States, [2]Department of Educational Leadership and Policy, College of Education (Courtesy appointment), The University of Texas at Austin, Austin, TX, United States, [3]Office of Academic Technology, Office of Academic Affairs, The University of Texas at Austin, Austin, TX, United States, [4]Department of Applied Physics, School of Engineering and Applied Sciences (Courtesy appointment), Harvard University, Cambridge, MA, United States, [5]Department of Psychology, College of Liberal Arts, The University of Texas at Austin, Austin, TX, United States

Generative AI presents opportunities and challenges for higher education stakeholders. While most campuses are encouraging the use of generative AI, frameworks for responsible integration and evidence-based implementation are still emerging. This Curriculum, Instruction, and Pedagogy article offers a use case of UT Austin's approach to this dilemma through an innovative generative AI teaching and learning chatbot platform called UT Sage. Based on the demonstrated benefits of chatbot technologies in education, we developed UT Sage as a generative AI platform that is both student- and faculty-facing. The platform has two distinct features, one a tutorbot interface for students and the other, an instructional design agent or builder bot designed to coach faculty to create custom tutors using the science of learning. We believe UT Sage offers a first-of-its-kind generative AI tool that supports responsible use and drives active, student-centered learning and evidence-based instructional design at scale. Our findings include an overview of early lessons learned and future implications derived from the development and pilot testing of a campus-wide tutorbot platform at a major research university. We provide a comprehensive report on a single pedagogical innovation rather than an empirical study on generative AI. Our findings are limited by the constraints of autoethnographic approaches (all authors were involved in the project) and user-testing research. The practical implications of this work include two frameworks, derived from autoethnographic analysis, that we used to guide the responsible and pedagogically efficacious implementation of generative AI tutorbots in higher education.

KEYWORDS

generative AI (GenAI), chatbots, responsible AI, instructional design (ID), educational technology, higher education, science of learning, teaching and learning

Introduction

Background

In the 1970s, inexpensive, hand-held calculators sparked a revolution in math education (Ellington, 2003; Raymond, 2024). After learning basic arithmetic, students could relegate tedious paper and pencil calculations to machines, opening up the opportunity to work on more interesting problems. Educators, however, faced a sea of ambiguity. Would students use

these tools to cheat? Would they lose computational skills by offloading too much to a piece of hardware? Could the calculator help advance student learning and solve long-standing problems, such as student motivation, in math education?

At present, the higher education discourse on generative AI parallels much of the early 1970s viewpoints on calculators (see Science News, 1975). Technically, generative AI and calculators represent radically different academic technologies. Lodge et al. (2023) emphasize that even though it is tempting (and popular) to do so, comparing the two oversimplifies the complexity of generative AI. For example, "generative AI could be described more as a technological infrastructure, like electricity, and not a single tool" (Lodge et al., 2023, para 4). That said, higher education faculty, administrators, and students today face a pedagogical dilemma analogous to the 1970s. Should we adopt generative AI without clear empirical evidence of how the tool might help, hinder, or harm student learning? How can we do so when so many unresolved questions about ethics, privacy, environmental impacts, bias, and career impacts relative to generative AI abound?

The existing situation: generative AI adoption and the teaching and learning landscape

Empirical research is a slow process, and so it can take years (or decades) to build up an evidence base about the efficacy of a new technology. Generative AI is not just "here" in that it is widely available throughout society, it is also solidly here and freely available on campuses worldwide. A study of 116 major research institutions in the United States found that most campuses are encouraging generative AI use (McDonald et al., 2025). Not only that, most of those same campuses also provide guidance to support generative AI adoption. Higher education leaders who are AI forward are aware of the importance of minimizing the digital divide and preparing students for a future where AI is ubiquitous. Students, moreover, want (and need) more than just access: They want generative AI lessons, especially concerning ethical adoption, incorporated into classroom learning (Cengage Group, 2024). Most faculty want to support student learning, but they may be unclear about how to do so with generative AI since it is so new. In addition, while some empirical studies correlate the use of generative AI with improved student learning outcomes (see Lo et al., 2025, Yilmaz, et al., 2023, Zhu, et al., 2025), generalizability and statistical effects vary widely. As such, institutions find themselves in a position whereby they need to lead their campuses toward the responsible adoption of generative AI in a rapidly shifting landscape of highly unresolved, high-stakes questions related to student learning.

While the impact of generative AI on student learning is evolving, general principles of responsible adoption of AI in teaching and learning do exist (U.S. Department of Education, Office of Educational Technology, 2023; WEF, 2024; McDonald et al., 2025). So too does firmly established, long-standing evidence of how students learn best (National Research Council, 2000; Ambrose et al., 2010; Hattie, 2015; National Academies of Sciences, Engineering, and Medicine, 2018). For example, drawing on the science of learning, it is clear that student learning is optimized when educators design their courses using student-centered, active learning approaches (Ambrose et al., 2010;

Schell and Butler, 2018). However, the large majority of higher education faculty are disciplinary specialists rather than pedagogical experts, so they may be unfamiliar with the scholarship of teaching and learning and how to apply it within an AI context. Moreover, faculty gaps in pedagogical knowledge may lead to inadvertent replication of teacher-centered designs in college classrooms.

Learning science research is both extensive and dense, which has led to a number of publications aiming to translate findings to practice (see National Academies of Sciences, Engineering, and Medicine, 2018; National Research Council, 2000). Improving one's teaching using principles from the science of learning takes time and effort, both of which are in short supply among research-active faculty. While information on how people learn best is plentiful, the realities of the faculty workload present a challenge for educators and institutional leaders who aim to advance the academic mission. Some institutions offer instructional design services to bridge these gaps.

With backgrounds in both learning theory and technology-enhanced pedagogy (Kumar and Ritzhaupt, 2017; Pollard and Kumar, 2022), instructional designers offer a valuable resource to faculty who want to build technological pedagogical content knowledge—or that special knowledge base for teaching specific content with technology (Voogt et al., 2013). Not all faculty are open to working with instructional designers, however (see Pollard and Kumar), and at major research universities, the need for quality instructional design consultation far exceeds available resources.

Advancing high-quality pedagogical practices by blending generative AI and learning science in a chatbot

The Office of Academic Technology Team at the University of Texas at Austin launched a generative AI development project to explore whether responsible adoption of emergent technology could help scale the use of learning science-driven instructional design at a major public research university. The purpose of this Curriculum, Instruction, and Pedagogy article is to offer a use case of an innovative generative AI chatbot designed from the ground up called UT Sage. For context, this paper focuses on the process of locally developing and alpha and beta testing an AI chatbot in higher education and is not an empirical study. We describe our conceptual approach to chatbot design and deployment, and detail two evidence-based frameworks that guided our design decisions. These frameworks represent replicable elements that higher education stakeholders can adapt to guide chatbot or other generative AI development efforts in their own instructional contexts.

UT Sage is a generative AI chatbot that is both student- and faculty-facing. AI chatbots are not new in education. In two separate meta-analyses covering AI chatbots, Okonkwo and Ade-Ibijola (2021) and Winkler and Soellner (2018) identified a host of potential benefits aligned with chatbot technology when educators deploy them for teaching and learning purposes, including student engagement, memory retention, access, metacognition, and self-regulation. Although these studies precede the influx of generative AI in education, established literature on AI chatbots in teaching and learning along with newewer works (see Lo et al., 2025, Yilmaz, et al., 2023, Zhu, et al., 2025), form a solid foundation from which to begin generative AI adoption initiatives on university campuses.

The UT Sage user experience for students is similar to other chat or tutorbot interfaces. Where UT Sage differs from other generative AI chatbot experiences is within its faculty-facing "builder bot" or custom-GPT features. Behind the scenes of the student-facing tutorbot, UT Sage functions as an always available, learning-science-driven, virtual instructional design coach or agent. The builder bot is a helper agent intentionally programmed to promote virtual instructional design coaching rooted in learning science research. With its dual nature feature of student tutorbot and instructional design agent, we believe UT Sage is a first-of-its kind application to integrate the science of learning with generative AI custom GPT technologies for classroom use.

This article begins with a broad overview of UT Sage as an educational innovation activity. We detail the key features that support the use case of UT Sage as a scalable, virtual instructional design agent. We include a methodology section to situate the project, while acknowledging the limitations of a non-experimental study. Then, we provide an overview of results from our assessments of UT Sage so far. Finally, we close with a discussion of the practical implications and lessons learned from our effort to scale learning-science driven instructional design coaching using a generative AI agent. After reading this case study, we expect higher education faculty and leaders to have an example for how to navigate the dilemma-laden landscape that broad, open-access to generative AI has brought to higher education. We offer two evidence-based frameworks we used to guide the local development of a generative AI chatbot. UT Sage serves as one early effort to adopt generative AI in higher education by integrating responsible AI and learning science principles with emergent technologies.

We want to be clear from the outset that our aim is not to replace or limit the role of instructional designers in higher-education institutions or to reduce faculty autonomy in course design. Teaching is an inherently human task, and what we offer through Sage is only a small part of what an instructional designer can do when engaging with faculty. Instead, the goal of this project is to improve teaching practice by scaling introductory elements of instructional design through the use of generative AI to bridge the gap between the supply of and demand for instructional resources on our campus. Without administrative intention and adherence to responsible AI principles, automation of course design will lead to deleterious effects on student, faculty, and designer roles. Automating basic elements of instructional design may also require designers and faculty to develop new competencies in the ethical and responsible implementation of generative AI in the classroom that aligns with the academic mission. When implemented with clear intention and responsible adoption principles, however, tutors like Sage may also open opportunities for instructional designers, technologists, and faculty to create innovative approaches to learning experiences that support transferable knowledge and skills.

Educational innovation activity: UT Sage

UT Sage overview

UT Sage is a platform that provides a scalable, virtual instructional design agent (the builder bot) to aid instructors in creating their own tutorbots for students. Our vision was to enable instructors to conceive of an idea for a student-facing chatbot tutor, have a conversation with the Sage agent to refine their vision, upload resources, and deploy their

TABLE 1 A list of the attributes that UT Sage uses to configure tutors for instructors and the related inquiries used by the instructional design agent as part of the conversational builder bot.

Tutor attribute	Instructional design coach inquiry
Topic	What topic would you like to create a tutor for today?
Learner Description	Who are your learners? Describe things like their likely academic year, majors, and minors. How large is their class and how is it delivered? What prior knowledge might they have about this topic, or what prior knowledge gaps might they experience. What is their motivation like for the topic?
Learning Outcomes	What are your learning outcomes for this tutor? What would you like your students to know about the topic? What would you like them to be able to achieve? What kinds of attitudes would you like them to gain or develop?
Topic Importance	Why is this topic important to learning in your class? Why is this topic important, generally, for students to learn?
Common Misconceptions	List common difficulties, misconceptions, inaccurate knowledge, or challenges that your students have with this topic. How have you helped your students work around these in the past?
Learning Activities	What kinds of activities would you like students to do when they engage with this tutor? For example, would you like them to quiz themselves or practice in some other way?
Training Documents	What kinds of resources would you like to upload to configure this tutor?

tutorbot to students in a few hours or less. As an agent, Sage is built to provide instructional design coaching with faculty to help them build effective tutorbots based on established learning science principles. Sage asks instructors the questions found in Table 1 to gather information about the learners and the desired learning experience. Once an instructor's tutorbot is created and shared, students can start a conversation with the tutor to supplement their knowledge of a topic. Tutorbots in Sage offer the experience of using chatbots to learn using generative AI tools, but with the assurance that the content knowledge loaded into those tools has been vetted by their instructors and adheres to the University's information security policies. Another unique aspect of Sage compared to other generative AI chatbots is that it is designed to operate at the topic or lesson plan level, rather than a full-course level. This decision was made to align Sage with a more typical tutor experience and to reduce the learning curve for a faculty member who may want to build a tutor bot.

UT Sage as AI-tutor and instructional design agent

The UT Sage platform is made up of two distinct elements:

(1) The builder bot instructor interface is where instructors can create tutors according to their own instructional needs. Instructors can chat with an instructional design

agent that asks them about what topic they'd like to build a tutorbot for, who their learners are, and how they'd like to define their learning outcomes as detailed in Table 1 and illustrated in Figure 1. The builder chatbot will make suggestions or pose questions to help guide the faculty in creating their tutor. In addition, the agent prompts instructors to document common misconceptions or difficulties students might have and any unique ways the faculty member has found for addressing those misconceptions. If instructors would like to adjust their tutor, they can also make changes to all of its parameters using a configuration form (see Figure 2). Additionally, faculty can upload and categorize three different text-based resources to train the tutor on the tutor topic. For example, a user can upload an administrative document, an assignment, or notes, and UT Sage will incorporate the information into conversations appropriately. For example, information parsed from assignment documents is handled with less literal transcription and more directed inquiry.

Content from administrative and notes documents is more directly integrated into tutor responses. As instructors build their tutors, they can also test the student experience in the Student Preview window on the right. The builder bot and training interface are illustrated in Figures 2,3.

(2) Students can access and use UT Sage tutors after their instructors have created, shared, and published them. The student-facing interface is illustrated in Figure 4. Tutorbot-student facing conversations are programmed to be helpful and to encourage students to engage in Socratic dialogue by asking questions at the end of appropriate interactions. Tutorbots use training documents uploaded by instructors as the first and best source of information. They do not engage in conversations about unrelated subjects. The tutor maintains a memory of what it has discussed with students previously, but a new session can also be created if students wish to start a new line of inquiry. A history of these conversations is maintained for students and accessible in the chat interface.
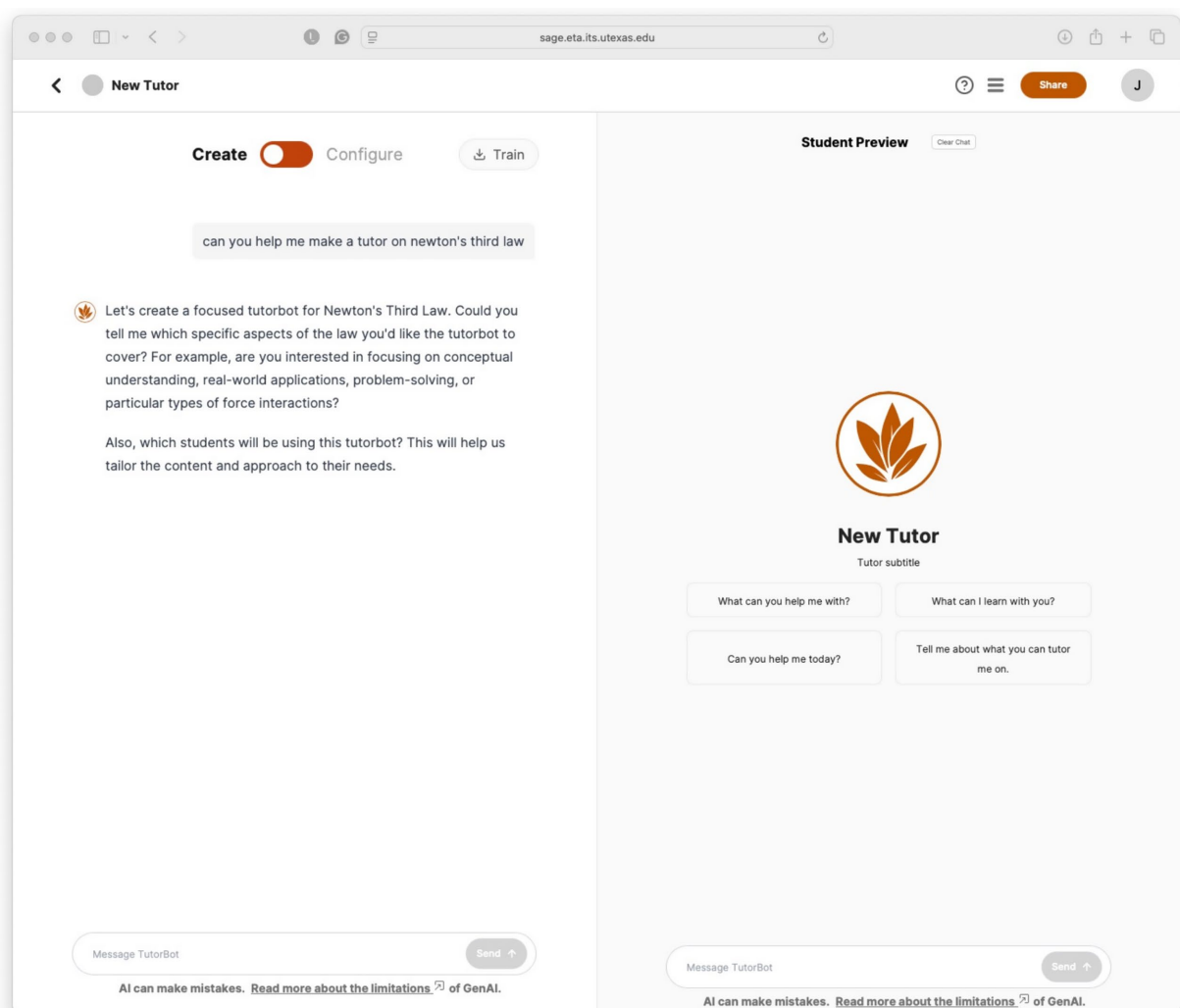

FIGURE 1
UT Sage's instructor-facing instructional design agent (Left) with student view test window (Right).

FIGURE 2
UT Sage's instructor-facing instructional design configuration form (Left) with student view test window (Right).

Each of these functions, builder bot and tutorbots, can be accessed via the platform homepage, which features all of the tutors that the user has access to. Students can see their tutors organized according to term, and instructors can edit or test any of their tutors from this page.

Figure 1 illustrates the instructor-facing experience with the instructional design agent on the left, with a preview window that instructors can use to test out the student-facing tutorbot they are building. Instructors use the configuration (Figure 2) and training (Figure 3) interfaces to refine and assess their tutors. The remainder of the configuration form includes the categories outlined in Table 1: learning outcomes, topic importance, common misconceptions and workarounds, learning activities, and "conversation starters" to help guide students who may not know how to begin. Figure 4 illustrates the student-facing experience with a tutorbot. In Figure 4, the instructor has created a tutorbot to help students with logistic regression. Students can get started with one of the conversation starters or type in their own text.

Technical details

Sage uses, at time of writing, the Claude 3.5 Haiku and Claude 3 Sonnet large language models (LLMs) to understand what students are asking and answer with context from topic-specific information using a retrieval-augmented generation (RAG) pipeline. Access to learn with Sage is free and available to students 24/7. Because this platform is owned by the University and students and faculty engage after logging in with their university ID, their input and output is protected by the University's highest data security and privacy standards.

Sage is a collaboration between the UT Austin's Office of Academic Technology and Enterprise Technology group, with the former offering product requirements and design and the latter developing institutional infrastructure, the user interface, and connecting underlying technologies. The prompts that power Sage's tutors were developed in partnership with AWS, which approached UT Austin about finding applications for generative AI technologies in higher education.
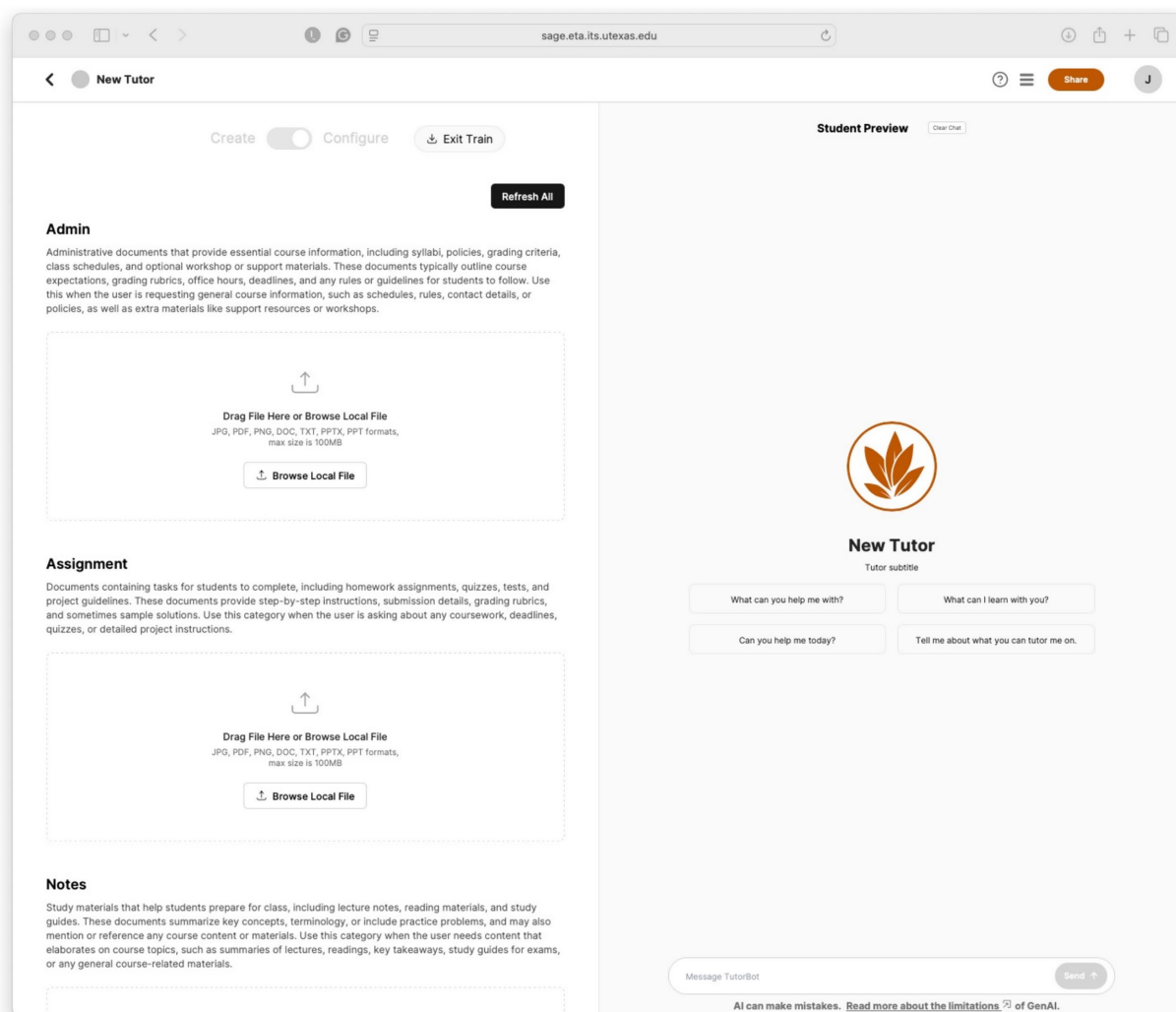
FIGURE 3
UT Sage's instructor-facing instructional design resource interface (Left) with student view test window (Right).

Learning environment

UT Austin is a large, public, R1 university with 19 colleges and schools. 51,913 individuals were enrolled as students in Fall 2023. Of those students, 56.3% are federally identified as women and 43.8% as men. 80.1% or 42,444 are undergraduate students, and 19.9% or 9,469 are seeking graduate degrees. These students are distributed among 156 undergraduate degrees and 237 graduate programs. 3,917 faculty were employed by the university for the 24–25 academic year and about 48.7% are tenure or tenure-track and 51.3% were professional or non-tenured (University of Texas at Austin, 2024).

Given the size of the student body and the breadth of available educational programs, the instructional needs and circumstances of these students and faculty are highly varied. A small handful of schools and departments have dedicated instructional designers, educational developers, and educational technologists on staff to address the needs of faculty, but the availability of these services across campus is inconsistent. While centralized

offices offering support for course design and technology implementation, such as the Office of Academic Technology and Center for Teaching and Learning, are available for consultation, the need for flexible access to personalized learning experience design advice has been recognized by central administrators.

Principles and frameworks underlying UT Sage

Responsible adoption of generative AI

The literature on the responsible adoption of generative AI in education—both in K12 and higher ed—calls for balancing its transformative potential of the new technology with active efforts to address its limitations and potential dangers (Saaida, 2023; WEF, 2024; McDonald et al., 2025). The UT Sage initiative involved a number of design decisions aimed at maintaining such balance. Prior to
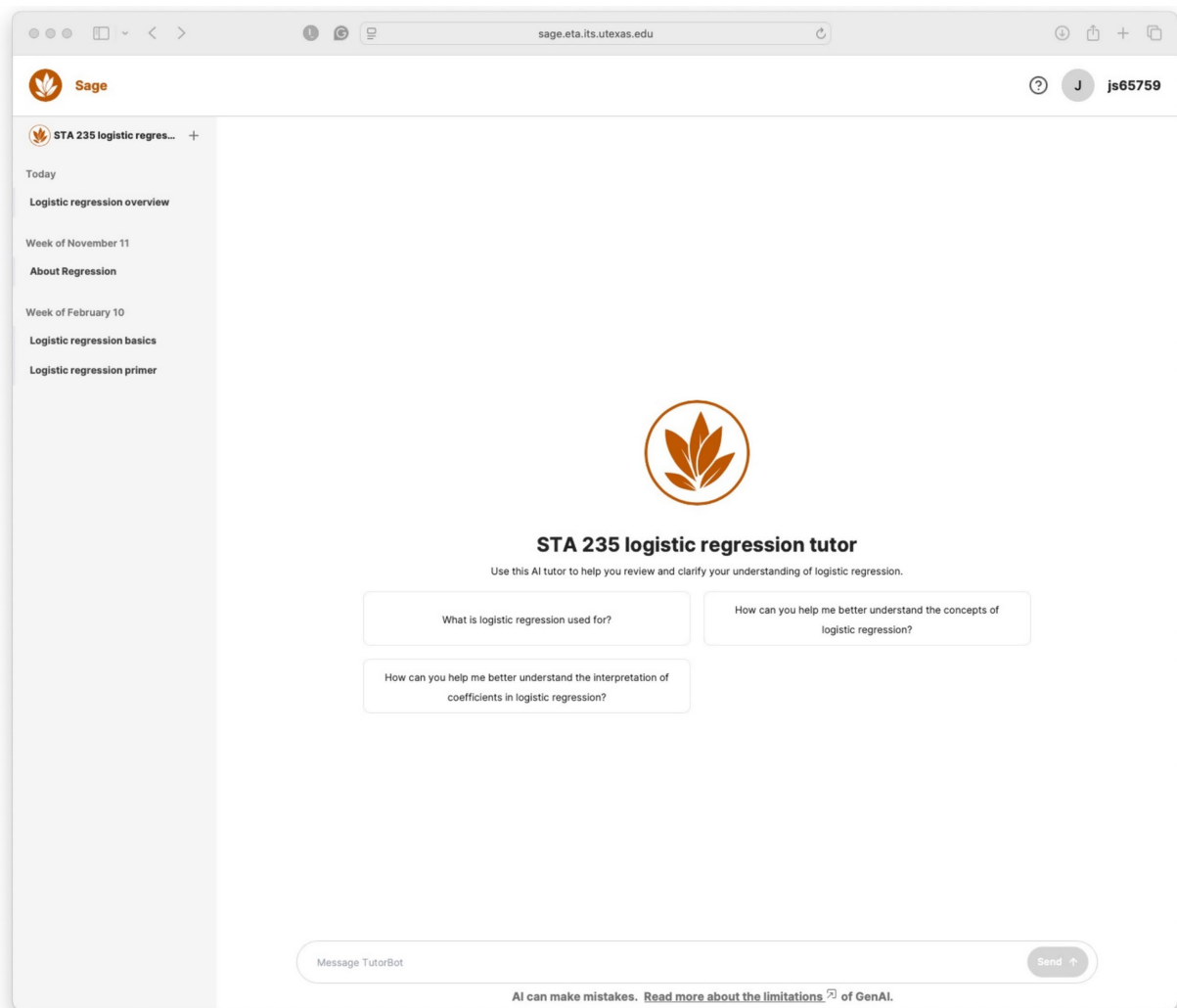
FIGURE 4
An example of a student-facing tutorbot chat interface in UT Sage called 'Statistics 235 logistic regression tutor'.

conceptualizing Sage, we developed the AI-Forward - AI-Responsible Framework (Office of Academic Technology, UT Austin, 2024) to guide campus to engage in responsible adoption of generative AI for academic use.

AI-Responsible/AI-Forward framework

Our AI-Responsible/AI-Forward framework calls for embracing generative AI for teaching and learning while also acknowledging that the technology also has significant limitations. The framework defines responsible use of generative AI tools for teaching and learning as using generative AI in ways that foster the achievement of learning outcomes and not using it in ways that would negate or inhibit the realization of those outcomes (Office of Academic Technology, UT Austin, 2022). We drew on the "human-in-the-loop" concept to develop this framework (U.S. Department of Education, Office of Educational Technology, 2023). Human-in-the-loop generative AI emphasizes that students and teachers must always be involved and have agency when it

comes to the adoption of AI tools. Our definition aims to empower educators to decide for themselves (1) how generative AI might improve student learning of specific topics and (2) to be transparent with students about why and how generative AI might help them achieve specific learning outcomes, or on the other hand, inhibit or harm their learning. We encourage faculty to foster a climate where students can become the architects of their own ethical frameworks in light of such transparency.

To help support AI literacy and bolster the responsible side of the balance needed for effective adoption, we also developed what we refer to as the "Big 6," which detail six limitations of using generative AI for learning in particular (Office of Academic Technology, UT Austin, 2024) as follows: Data privacy and security, hallucinations, misalignment, bias, ethics, and cognitive offloading. The limitations of generative AI become even more complex at scale. Efforts to adopt generative AI across contexts require higher education leaders to engage in consistent grappling with issues such as the digital divide, training and algorithmic biases, risks of exposing student data, and over-reliance on AI tools in ways that short-circuit the academic honor code and productive struggle (Bjork and Bjork, 2020).

While the University now provides enterprise-level access to Microsoft Copilot, at the time we began developing Sage, the campus did not have an open-access, approved generative AI tool for educational use. We used the AI-Responsible/AI-Forward framework to determine a set of four design strategies and related design principles to build Sage highlighted in Table 2.

This documentation provides an overview of principles of responsible AI that we used to guide the need to balance embracing new and rapid diffusion of a new technology in teaching and learning, with the need to ensure transparency and education related to its hazards. Institutional leaders, faculty, and other stakeholders can use or adapt these principles to help guide their responsible AI efforts.

The Tetrahedral Model of Classroom Learning

Educational technology scholars emphasize that the killer app feature inherent in an AI chatbot is tied to such tools' abilities to personalize or customize student learning experiences (Bii, 2013; Winkler and Soellner, 2018). We adopted this perspective by conceptualizing UT Sage as an AI tutorbot that could be trained by faculty through an instructional design agent programmed specifically to elicit an educator's pedagogical content knowledge (PCK; Shulman, 1986). PCK is a special blend of disciplinary expertise and depth of understanding around how students best learn content within a discipline. Faculty build PCK throughout their careers and develop an intuition for what makes learning a particular topic difficult and how to help students overcome those challenges. Because it is complex knowledge (Shing et al., 2015), PCK is often deeply internalized, but not externalized in one's teaching practice beyond typical artifacts, such as a syllabus. UT Sage was conceptualized to allow educators to capture intuitions like this and document them through custom training a tutorbot using their own PCK.

Principles of learning

How do students learn best? One answer to this question is that students learn best when educators design learning experiences that center on the learner and their needs relative to the content (Ambrose et al., 2010; Hattie, 2015; National Academies of Sciences, Engineering, and Medicine, 2018; Schell and Butler, 2018). Learner-centered approaches contrast with topic-centered or instructor-centered teaching, where delivering the content alone is the central point of focus. Learner-centered teaching is generally guided by PCK, where topic-centered teaching often bifurcates content and pedagogy. While learner-centered teaching has caught on in some sectors of higher education and empirical evidence supports its use (Shing et al., 2015), it remains that most faculty are trained to be disciplinary versus pedagogical experts, and as such, their teaching approaches replicate the topic-centered instruction they themselves received. Learning-science-trained instructional designers are aware of the benefits of learner-centered teaching and can help instructors transition their approaches. The problem we worked to address with UT Sage is supply versus need for instructional design at a R1 campus.

Moreover, Chen et al. (2025) recently documented that, while generative AI provides support for teachers in building lesson plans, AI-generated content predominantly promotes teacher-centered approaches, "with limited opportunities for student choice,

goal-setting, and meaningful dialogue" (p1). Ensuring generative AI's promise for teaching and learning requires leaders to intentionally guard against building systems or chatbots that replicate ineffective teaching. Chen et al. also demonstrated how appropriate prompt engineering can help mitigate inherent teacher-centered biases in generative AI.

Sage was designed from the ground up to drive student-centered tutoring with a generative AI chatbot. The Tetrahedral Model of Classroom Learning (TMCL) (Schell and Butler, 2018) depicted in Figure 5 is a student-centered model that highlights four key components that any educator must consider to facilitate effective learning in their classroom. We used these four components to define a set of additional design strategies and principles to help faculty train their tutorbots in Sage. It is worth noting that instructional design is an established field that cannot nor should be replaced by a tool like Sage. Teaching is an inherently human task, and what we offer through Sage only touches the surface of what can and should be accomplished through a strong instructional design relationship. We hope that by initiating ways to surface and interact with one's own PCK, we will help promote effective lesson plan design to those who do not practice or are not aware of learner-centered teaching and spark interest in developing deeper learner-centered teaching practices.

We designed the Sage's instructional design agent depicted in Figures 1–3 above to align directly with TMCL principles. For example, the scholarship of teaching and learning has established that prior knowledge strongly influences new learning (Ambrose et al., 2010; Hattie, 2015). This literature informed our decision to require instructors to document students' prior knowledge gaps during bot configuration. Similarly, Sage will coach a faculty member through the development of learning outcomes, which reflects longstanding research that demonstrates student achievement is correlated with clearly articulated goals and expectations. Finally, self-regulated learning theories (Ambrose et al., 2010; National Academies of Sciences, Engineering, and Medicine, 2018; Schell and Butler, 2018) bolstered our efforts to ensure the tutor posed questions to spark metacognition (the act of thinking about or assessing one's learning state).

Table 3 outlines each of the key learning science principles we used and how those principles were built into design requirements for Sage.

In summary, by carefully conversing with the instructional design agent within UT Sage (i.e., the builderbot), we designed and implemented a novel way for instructors to (1) begin engaging in learner-centered design following established principles; (2) customize their students' learning experiences with generative AI based on their own individual PCK in ways that are only possible through generative AI; and (3) surface, interact with, and incorporate their own PCK into customized, generative AI tutor bots for their students.
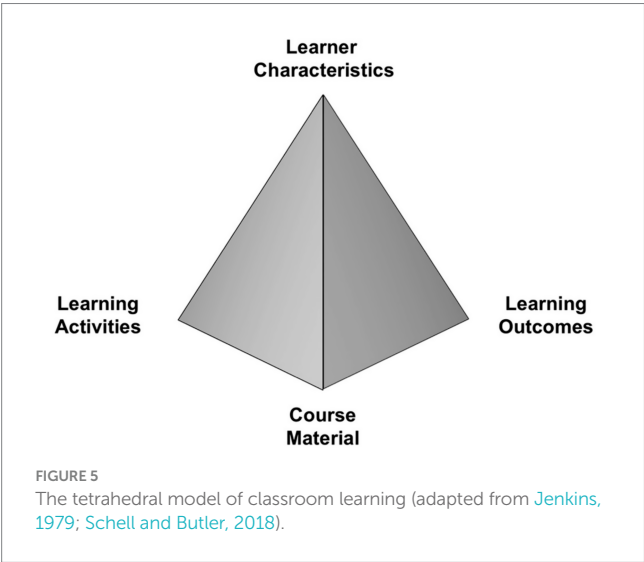
Methods

UT Sage pilot release life cycle and sampling

The primary purpose of this project was to develop software. As such, methodologically, we followed a standard, user-centered software lifecycle approach to developing, releasing, testing, and refining UT Sage with evaluation measures, data collection, and participant selection procedures that aligned with our production

TABLE 2 Design principles framework for responsible adoption of generative AI, illustrating the strategies and design principles used to build UT Sage to ensure responsible AI adoption.

Strategy	Principle	Description
Provide all responsible access to generative AI for teaching and learning	Equitable Access	To ensure equitable access to University resources to all learners, UT Sage was designed to be freely accessible to all faculty and students.
Offer generative AI tools that ensure information security resources are protected and accessible	Data Privacy and Security	Since UT Sage is designed to create and collect educational records, Data Security and Privacy were driving concerns. Moreover, students at The University of Texas at Austin maintain the intellectual property rights of materials they create or produce in their coursework. UT Sage was designed to align directly with the University's Information Resources and Security Resources and to provide data security, intellectual property and FERPA protections. In addition, our team is partnering with the University's Digital Accessibility Center to follow best practices and accessibility requirements.
Infuse learning science into the adoption of generative AI	Focus on prior knowledge and learning outcomes	To configure a tutor on any topic Sage, instructors must think about their students' characteristics, including the state of their prior knowledge acquisition on the topic. They must also document their learning outcomes or what they hope students will know, be able to do, or the attitudes they might develop as a result of using the tutor.
Practice transparency and support AI literacy	Balance	While using UT Sage, Students see an always-on display above the chat interface (see bottom of Figure 2) that reads: "AI can make mistakes. Read more about the limitations of Gen AI" which directs to our Big 6 limitations page (Office of Academic Technology, UT Austin, 2024): Data security and privacy, hallucinations, misalignment, bias, ethics, and cognitive offloading.

FIGURE 5
The tetrahedral model of classroom learning (adapted from Jenkins, 1979; Schell and Butler, 2018).

goals. We designed the project to align with the following phases: pre-alpha, alpha, closed-beta, open-beta, and general availability. For the purposes of this article, we employed authoethnographic methods by systematically analyzing and describing a teaching and learning innovation that all three authors were involved in (see the Acknowledgements section). Below we provide details on pilot participant sampling and limitations, data analysis, and each phase of the pilot implementation.

Pilot participant selection

For the pre-alpha through the closed-beta phases of the project, faculty participants were recruited using convenience sampling via University-wide announcements and programming events. During the open-beta phase, both convenience and snowball sampling – where faculty heard about UT Sage from other users, were employed.

Student participants were recruited through convenience sampling and limited by their enrollment in courses taught by the faculty participating in the pilot. The first author participated in the Alpha testing with students to assess the alignment of the tutor with the original concept. The second and third authors participated pre-alpha through beta testing with the builder bot.

An important limitation of our alpha and beta testing was that we prioritized convenience sampling for the purposes of eliciting feedback on bugs, functionality, and general user experience. User-centered software development can prioritize the needs of immediate user preferences and may lead to solutions that are biased and do not generalize well across all users. The open-beta phase will address some of these risks by broadening participation beyond a convenience sample to the full instructor and student population at the University. This larger sample should enable more differentiated feedback that will better reflect a fuller range of user needs and contexts.

Pilot data collection and analysis

Using an issue tracking process, the Sage team collected quantitative and qualitative data documented from surveys, narrative feedback, and observational feedback in each phase of testing. Data was thematically coded as a bug or as a feature enhancement and translated into design requirements.

Pre-alpha testing proof of concept

We began developing a proof of concept for the vision of UT Sage as an instructional design agent and student-facing tutorbot in the Summer of 2024. During the pre-alpha phase, we wrote narrative scripts for how the instructional design agent should interact and function with the users, as well as created wireframes for the interface. As is standard practice, pre-alpha iteration was completed internally with key stakeholders and project team members only.

TABLE 3 Design principles framework for learning science-driven adoption of generative AI, illustrating the strategies and design principles used to build Sage to ensure established learning science drove the generative AI tutorbot experience.

Strategy	Principle	Description
Consider your learner characteristics and how they might influence their learning experiences	Learner Characteristics	There few influences that have more power to determine student learning than their specific learning characteristics, especially their prior knowledge and previous exposure to the topics (Ambrose et al., 2010). Learning is influenced when teachers based their teaching on "what students bring to the subject" (Hattie, 2015, p. 81). As such, one of the first requirements for building a bot within UT Sage is to document learners' characteristics relative to the course context and topic.
Clearly articulate learning outcomes	Learning Outcomes	Student learning outcomes are the things students should know, be able to do, and the attitudes they should hold after completing a learning experience (Allan, 1996; Wiggins and McTighe, 2005; Tyler, 1949; Schell and Butler, 2018). Large meta analyses focused on higher education indicate that when educators clearly articulate learning outcomes, student learning is heightened (Hattie, 2015). One reason for this finding is that when students can identify what success might look like they can more easily self-regulate their learning to achieve those outcomes. Moreover, with clearly articulate learning outcomes instructors can more easily evaluate their impact on learning (see Hattie). However, effective learning outcomes can be both elusive and difficult to develop, especially for higher education who have had limited or no pedagogical training. UT Sage provides an interface for instructors to develop effective learning outcomes to guide their tutor bots.
Develop learning activities that promote active versus passive learning	Active Learning	Active learning is variously defined in the literature; we define active learning as acquiring knowledge, skills or attitudes through intentionally self-directing one's learning activities and constructing rather than "receiving" content knowledge (Schell and Butler, 2018). UT Sage was conceptualized to support instructors training tutors to promote active learning heavy hitters such as metacognition, retrieval-enhanced-learning, and corrective as well as evaluative feedback. For example, instructors can configure their bots to nudge students to start off a session by quizzing themselves or analyzing a piece of text. In addition, there is a dedicated section in the bot-builder to encourage instructors to describe these activities but as of the publication date, a technical barrier has prevented implementation.
Optimize faculty time by leveraging course materials you have pre-built and found effective	Course Materials	Most instructors have spent extensive time developing and curating course artifacts to support student learning, including readings, documents, slides, images, videos, audio and more. Knowing when and how to deploy such artifacts is a key component of PCK. Our Beta version of Sage supports uploading of text-based documents that the technology then incorporates into training tutorbots. If additional releases of Sage are deployed, we expect additional media types to be included in future versions.

Alpha testing

Alpha testing was staged in early Fall of 2024 and included testing of the proof of concept and internal functioning with minimal features and known errors. Issue tracking was implemented at this stage. The team was tasked with creating a user interface and engineering a prompt that integrated the learning principles identified in the initial design process with the LLM and RAG pipeline.

Among the fifteen faculty enrolled in the closed-beta testing phase, which focused on the creation and tuning of tutors, five colleges (College of Natural Sciences, College of Education, McCombs School of Business, College of Fine Arts, and the College of Liberal Arts) and 11 fields and disciplines (chemistry, statistics, computer architecture, information studies, information management, business management, entrepreneurship, marketing, design, higher education leadership, classics) were represented. A total of six of the tutors proposed were created and tested by instructors and their colleagues as part of the phase 1 alpha. Of these, two were shared with students for testing. One tutor was provided to a group of fifteen graduate students during a face-to-face class. The other was provided as a resource to a class of sixty undergraduate students for use during preparations for the course final exam.

Once the student-facing tutorbot interface (Figure 4) was functional, faculty worked with a human instructional designer specializing in AI (this paper's second author) to provide specifications for their tutors in a design document similar to what one might use as part of an instructional design consultation and following the TMCL in Figure 5 above. Sage used faculty responses to the prompts in Table 1 to define and create six tutors for closed-beta testing. These

tutors addressed varied pedagogical needs in diverse fields of study. Examples of tutors conceptualized and created by instructors and the Sage team include the following cases.

- Case 1: A tutor focused on aiding undergraduate business students in a Statistics course in understanding concepts related to logistic regression. Resources were provided to train the tutor to advise students on how to determine when to use logistic versus linear regression and their underlying mathematical distinctions.
- Case 2: A tutor designed to coach senior-level chemistry majors in the application of analytical chemistry techniques. The tutorbot was designed as a study aid and bridging activity for students who are learning concepts in their lecture-based instruction and performing them in the lab.
- Case 3: A tutor whose primary purpose is to coach graduate students in design and education in the creation of learning outcomes. Depending on their background, these students might have congruent gaps in knowledge in design and learning theory, respectively. This tutor can evaluate outcomes provided by the student and advise them on improvements using the resource Bloom's Taxonomy.

Alpha testing results

The purpose of the Alpha user testing was to get initial feedback on the usability and perceptions of the chatbot. Data collection methods included two surveys (included in the Supplemental materials) and an option for faculty and students to give open, narrative feedback

via e-mail, and one, autoethnographic live observation conducted by the first author.

Faculty feedback

The faculty reported an overall positive experience using their tutors and unanimously agreed that it could aid students in meeting their stated learning outcomes, however, we did not test this perception. They also noted that the information provided was accurate and the answers were clear. They also provided suggestions for interface features (such as removing in-text citations and automatically naming chat sessions) and changes to the way the tutor interacts with students. Specifically, they requested that the length of responses be reduced; that the tutors determine when it should use Socractic questioning to engage students with topic concepts; and to avoid being apologetic when it could not retrieve additional information for the user.

Student feedback

Of the seventy five users that were given access to beta tutors, 14 provided feedback in live observations and surveys. Student reactions to the tutors were mixed with many experiencing authentication, display, and other technical errors. Most acknowledged that the tutorbot helped them learn the topic at hand and met or exceeded their expectations for such a tool. Some also noted lengthy responses and numerous questions that the faculty had also pointed out to the team. One user provided in-depth feedback about the lack of customization in tutorbot responses for students who have reading disorders and other needs related to processing text, suggesting that they be able to have text read to them by the tool or adjust response output to their particular needs.

Many of the suggestions made during phase 1 alpha testing were implemented for the phase 2 beta and integrated into the interface shown in Figures 1–3.

Closed beta

Phase 2 of testing began in January of 2025 and was structured as a closed beta with a pool of invited testers of more than 40. In this phase, the instructor-facing builder tools were partially available with instructors being granted the ability to configure tutors through a form and to upload text resources to be ingested into the tutor's knowledge base. Additionally, tutors were shareable with anyone within the University or assigned to existing course rosters, so that student testing could be expanded. As of this writing, faculty can train tutors using the interface illustrated in Figures 1–3 instead of working directly with the designer.

Open beta

The next phase of testing is an open-beta where any staff or faculty member with an active University ID can designed a tutor and share it with their students. Key milestones for this phase include the addition of the following features.

- Conversational configuration where faculty can create new tutors by having a two-way conversation with the agent versus configuring the form in Table 1 and Figure 2. In addition to enabling an organic design experience, the agent will make suggestions about how to effectively tune and scope the tutor based on the learning characteristics and learning outcomes that the instructor has identified;
- Summary of student insights about common student questions and misconceptions about a topic. Sage will produce output for instructors to use for just-in-time teaching based on analysis of common student questions, misconceptions or other input and output.
- Integration of more input and output data processing tools that allow for Sage to ingest and properly respond with images, LaTeX, formatted code, and audio.
- Outcomes research planning to organize the assessment of the Sage platform across disciplines and implementations. Institutional Review Board processed studies will examine questions related to the effect of generative tutors on student learning outcomes, how to effectively train, test, and introduce tutors into course design, and student attitudes toward instructor-trained course topic tutors. Methodologies will be chosen to best fit each question, course, and field of study.

Along with these new features, we will continue to expand the scope of our testing, making use of the influx of data that new users will provide.

Discussion and implications

Our experience testing UT Sage has supported our motivations for developing a tutor-based chatbot, while also providing us with important feedback about how to improve platforms of this type in the future. Our aim was to provide a learning technology platform that leads faculty through the process of identifying the core elements of the tetrahedral model of classroom learning (i.e., learner characteristics, learning outcomes, learning activities, and course materials) using a conversational interface that would be comfortable for faculty to engage with ease. In this way, a simple instructional design task can be automated and we can mitigate the teacher-centered biases that may be inherent in current generative AI platforms (see Chen et al., 2025).

With the development of the builderbot, we were able to validate that a learner-centered process can be implemented in a way that supports student engagement across a variety of topic areas and levels of student expertise. Once implemented, the platform is relatively easy for faculty to use, so that they can quickly answer the instructional design questions and construct a bot for their course.

In addition to embedding principles of learner-centered design in the tutorbots, UT Sage has the benefit that it is always available to students, thereby increasing the amount of time that a proxy for the instructor can be accessed. Students frequently get stuck when reading complex material or working on difficult assignments at times when instructors and teaching assistants are not available. The tutorbot enables students to continue working on potentially frustrating assignments at a time convenient for them rather than just when human instructors are available.

That said, there are challenges that we have encountered as well. A human instructor can often sense levels of student frustration and can calibrate the degree to which they can lead students through a Socratic dialog when the student is asking for the answer to a question they are struggling with. The tutorbot is not sensitive to these aspects of student motivation, and so it may provide answers that are too long and may engage students in dialog longer than the student is comfortable with.

Planned enhancements to Sage include summaries of common themes and misconceptions that instructors can use to enhance direct instructional efforts. When instructors have insight into what tutors are helping students with, they can further refine learning outcomes for class sessions. In a similar fashion, information about what kinds of topics and learning activities are being selected for tutors by instructors can give instructional support staff in departments and colleges more insight into learning challenges.

In this way, we hope that UT Sage ultimately increases engagement between faculty, good instructional design, and instructional designers on campus. At present, many faculty do not have a deep understanding of the benefits of working with an instructional designer. By highlighting the instructional design capacities baked into the design of the builderbot, we give faculty a chance to get a first experience with instructional design and effective pedagogy. We hope that positive experiences with UT Sage increases faculty interest and willingness to work with generative AI and instructional designers to further improve their courses using evidence-based practices. These efforts may lead to additional ideas for builderbots to solve frequently encountered education problems in our courses.

While we did not empirically evaluate the relationship between UT Sage and the achievement of learning outcomes, we believe it is the most important direction for future research and practice in line with recent scholarship on the topic (Lo et al., 2025; Yilmaz and Yilmaz, 2023; Zhu et al., 2025). Research questions for future study of Sage include but are not limited to: How does the use of the learner-centered UT Sage tutor relate to student performance on assessments? What is the relationship between student self-efficacy on specific topics and use of UT Sage tutors tailored to those topics? How does performance on assessments or self-efficacy differ when we compare UT Sage with other generative AI tools that may have teacher-centered biases? In addition, we expect to explore a research agenda related to the adoption of generative AI by designing studies that investigate the relationship between the use of the UT Sage tutor and faculty self-efficacy with using generative AI and/or science of learning principles.

Finally, this article provides two frameworks in Tables 2, 3 to guide structured approaches to responsible adoption of generative AI in higher education. Specifically, higher education leaders can apply the design strategies and principles offered in this case study to integrate generative AI tools into teaching and learning in ways that are secure, pedagogically effective, responsible, transparent, accessible, and support AI literacy.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements.

Author contributions

Funding

Acknowledgments

at Austin, including Ladd Hanson, Eric Weigel, Alex Knox, Margaret Spangenberg, Mario Guerra, Jingwen Zhang, and Ahir Chatterjee.

Conflict of interest

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2025.1604934/full#supplementary-material

References

Allan, J. (1996). Learning outcomes in higher education. *Studies in Higher Education* 21, 93–108. doi: 10.1080/03075079612331381487

Ambrose, S. A., Bridges, M. W., DiPietro, M., Lovett, M. C., Norman, M. K., and Mayer, R. E. (2010). How learning works: Seven research-based principles for smart teaching. Newark, NJ: John Wiley & Sons, Incorporated.

Bii, P. (2013). Chatbot technology: a possible means of unlocking student potential to learn how to learn. *Educ. Res.* 4, 218–221.

Bjork, R. A., and Bjork, E. L. (2020). Desirable difficulties in theory and practice. *J. Appl. Res. Mem. Cogn.* 9:475. doi: 10.1016/j.jarmac.2020.09.003

Cengage Group (2024). 2024 graduate employability report: preparing students for the GenAI-driven workplace. Available online at: https://cengage.widen.net/s/bmjxxjx9mm/cg-2024-employability-survey-report (Accessed January 24, 2025).

Chen, B., Cheng, J., Wang, C., and Leung, V. (2025). Pedagogical biases in AI-powered educational tools: the case of lesson plan generators. *Soc. Innov. J.* 30, 1–8.

Ellington, A. J. (2003). A meta-analysis of the effects of calculators on students' achievement and attitude levels in precollege mathematics classes. *J. Res. Math. Educ.* 34, 433–463. doi: 10.2307/30034795

Ellis, C., Adams, T. E., and Bochner, A. P. (2011). Autoethnography: an overview. *Hist. Soc. Res.* 36, 273–290.

Hattie, J. (2015). The applicability of visible learning to higher education. *Scholarsh. Teach. Learn. Psychol.* 1, 79–91. doi: 10.1037/stl0000021

Jenkins, J. J. (1979). Four points to remember: a tetrahedral model of memory experiments, in *Levels of Processing in Human Memory (PLE: Memory)*, Eds. Laird S. Cermak and Fergus I. M. Craik (Hillside, NJ: Psychology Press), 429–446.

Kumar, S., and Ritzhaupt, A. (2017). What do instructional designers in higher education really do? *Int. J. E-Learn.* 16, 371–393.

Lo, N., Wong, A., and Chan, S. (2025). The impact of generative AI on essay revisions and student engagement. *Comput. Educ. Open*:100249. doi: 10.1016/j.caeo.2025.100249

Lodge, J. M., Yang, S., Furze, L., and Dawson, P. (2023). It's not like a calculator, so what is the relationship between learners and generative artificial intelligence? - UQ eSpace. *Learn. Res. Pract.* 9, 117–124. doi: 10.1080/23735082.2023.2261106

McDonald, N., Johri, A., Ali, A., and Collier, A. H. (2025). Generative artificial intelligence in higher education: evidence from an analysis of institutional policies and guidelines. *Comput. Hum. Behav. Artif. Humans* 3, 1–11. doi: 10.1016/j.chbah.2025.100121

National Academies of Sciences, Engineering, and Medicine (2018). How people learn II: Learners, contexts, and cultures. Washington, D.C.: The National Academies Press.

National Research Council (2000). How people learn: Brain, mind, experience, and school: Expanded edition. Washington, D.C.: National Academies Press.

Office of Academic Technology, UT Austin. (2022). Developing learning outcomes for your course. Available online at: https://provost.utexas.edu/the-office/academic-affairs/developing-learning-outcomes/ (Accessed January 31, 2025).

Office of Academic Technology, UT Austin. (2024). Addressing the limitations of using generative AI for learning. Available online at: https://provost.utexas.edu/the-office/academic-affairs/office-of-academic-technology/limitations/ (Accessed January 24, 2025).

Okonkwo, C. W., and Ade-Ibijola, A. (2021). Chatbots applications in education: a systematic review. *Comput. Educ. Artif. Intell.* 2, 1–10. doi: 10.1016/j.caeai.2021.100033

Pollard, R., and Kumar, S. (2022). Instructional designers in higher education: roles, challenges, and supports. *J. Appl. Instr. Des.* 11, 7–25. doi: 10.59668/354.5896

Raymond, C. (2024). AI and the case for project-based teaching. Chron. High. Educ. Available online at: https://www.chronicle.com/article/ai-and-the-case-for-project-based-teaching (Accessed January 24, 2025).

Saaida, M. B. E. (2023). AI-driven transformations in higher education: opportunities and challenges. *Int. J. Educ. Res. Stud.* 5, 29–36.

Schell, J. A., and Butler, A. C. (2018). Insights from the science of learning can inform evidence-based implementation of peer instruction. *Front. Educ.* 3, 1–14. doi: 10.3389/feduc.2018.00033

Science News (1975). Math in the schools: what's wrong? *Sci. News* 108:325. doi: 10.2307/3960361

Shing, C. L., Saat, R. M., and Loke, S. H. (2015). The knowledge of teaching – pedagogical content knowledge (PCK). *Malays. Online J. Educ. Sci.* 3, 40–55.

Shulman, L. S. (1986). Those who understand: knowledge growth in teaching. *Educ. Res.* 15, 4–14. doi: 10.3102/0013189X015002004

U.S. Department of Education, Office of Educational Technology (2023). Artificial Intelligence and the Future of Teaching and Learning. Washington, D.C. Available online at: https://tech.ed.gov/ai-future-of-teaching-and-learning/

University of Texas at Austin. (2024). Facts & figures | University of Texas at Austin. Available online at: https://www.utexas.edu/about/facts-and-figures (Accessed January 29, 2025).

Tyler, R. W. (1949). *Basic Principles of Curriculum and Instruction., ed. P. S. Hlebowitsh. Chicago, IL: University of Chicago Press. Available at:* https://press.uchicago.edu/ucp/books/book/chicago/B/bo17239506.html (Accessed January 29, 2025).

Voogt, J., Fisser, P., Pareja Roblin, N., Tondeur, J., and van Braak, J. (2013). Technological pedagogical content knowledge – a review of the literature. *J. Comput. Assist. Learn.* 29, 109–121. doi: 10.1111/j.1365-2729.2012.00487.x

WEF (2024). Shaping the future of learning: the role of AI in education 4.0. World economic forum. Available online at: https://www.weforum.org/publications/shaping-the-future-of-learning-the-role-of-ai-in-education-4-0/ (Accessed January 24, 2025).

Wiggins, G. P., and McTighe, J. (2005). *Understanding by Design.*, 2nd Edn. Alexandria, VA: Association for Supervision and Curriculum Development.

Winkler, R., and Soellner, M. (2018). Unleashing the potential of chatbots in education: a state-of-the-art analysis. *Acad. Manage. Proc.* 2018. doi: 10.5465/AMBPP.2018.15903abstract

Yilmaz, R., and Yilmaz, F. G. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Comput. Educ. Artif. Intell.* 4:100147. doi: 10.1016/j.caeai.2023.100147

Zhu, Y., Liu, Q., and Zhao, L. (2025). Exploring the impact of generative artificial intelligence on students' learning outcomes: a meta-analysis. *Educ. Inf. Technol.* doi: 10.1007/s10639-025-13420-z