Design, Implementation, and Early Engagement Outcomes of a Student-Led AI Tutor in Higher Education

Ahmed Alsharif, Utah Valley University, United States

The IAFOR International Conference on Education in Hawaii 2026
Official Conference Proceedings

Abstract

This paper presents the design, implementation, and assessment of a student-led Artificial Intelligence (AI) tutoring project at Utah Valley University (UVU). The system, integrated within the Canvas Learning Management System, was created to provide real-time, personalized academic support while upholding educational integrity. Its design intentionally modeled the scaffolding practices of human tutors, guiding students through problem-solving rather than supplying direct answers. Development was carried out by a small team of undergraduate students in collaboration with faculty mentors, supported by institutional grants and technology resources. The result was a mobile-friendly, scalable tool capable of supporting multiple courses. A preliminary assessment was carried out during the pilot deployment, drawing on surveys, analytics, and student reflections. Findings showed high engagement, with students reporting that the system enhanced understanding of course content and encouraged active learning. Technical evaluation confirmed reliable performance, and the project demonstrates that student-driven innovations can create meaningful educational technologies that promote accessibility, engagement, and equity in higher education. The findings also highlight both the pedagogical benefits and implementation challenges of introducing AI into authentic classroom contexts, offering a roadmap for institutions seeking to integrate AI responsibly into teaching and learning.

Keywords: generative artificial intelligence, AI tutoring, student engagement, higher education, large language models

Introduction

Generative AI is now a routine feature of students' academic environment, offering rapid access to explanations, examples, and iterative feedback. Evidence syntheses indicate that generative AI can improve learning performance and learner perceptions on average, but with substantial variability by context, learning design, and the role assigned to the AI tool. This variability implies that "AI in education" does not operate as a uniform intervention; outcomes are design-dependent and sensitive to how the AI is positioned in the learning process (Kasneci et al., 2023; Wang & Fan, 2025).

At the same time, rigorous empirical work cautions that unstructured AI support can undermine durable skill acquisition if students use the tool as a shortcut during practice. A large-scale field experiment in mathematics showed that unfettered access to a "ChatGPT-like" interface improved practice performance but reduced post-access independent performance; safeguards that shift the system toward tutorial hinting rather than answer-giving mitigated these effects (Bastani et al., 2025). This evidence is particularly consequential for higher education, where learning goals emphasize transferable competence and where educational integrity and assessment validity are core institutional responsibilities (Cotton et al., 2024; Kofinas et al., 2025).

The present study reports a student-led AI Tutor integrated into Canvas Learning Management System (LMS). The focus of this paper is deliberately narrow: it emphasizes early engagement outcomes, pedagogical implications, and adoption/implementation lessons, rather than detailed empirical replication of prior reporting. Engagement is treated as a necessary (but not sufficient) condition for learning improvement: if students do not adopt a tutoring system, the system cannot influence persistence, time-on-task, or constructive problem-solving behaviors. Where adoption occurs, the critical design problem is ensuring engagement is *cognitively productive*, not merely frequent use (Chi & Wylie, 2014; Kahu, 2013).

Unlike faculty- or vendor-led AI deployments, this initiative positioned undergraduate students as primary system builders, making the development process itself a form of high-impact pedagogy.

Foundations and Related Work

This project's instructional design rests on the idea that "AI tutoring" should behave like tutoring rather than answer generation. The theoretical basis begins with Vygotsky's (1978) concept of the Zone of Proximal Development which frames learning as what a learner can achieve with appropriate support beyond independent capability. Building on this foundation, Wood et al. (1976) introduced the term "scaffolding" to describe the tutorial supports that enable learners to complete tasks otherwise beyond them, with support that can be withdrawn as competence increases.

Within educational technology research, scaffolding has been operationalized as design patterns that structure learner activity, make thinking visible, and shape inquiry through prompts, templates, and just-in-time feedback. In particular, Quintana et al. (2004) offer a design framework for scaffolded software in science inquiry contexts, emphasizing scaffolds that align with inquiry practices and the challenges learners encounter during those practices. For AI tutoring in an LMS, this body of work supports an important design stance: scaffolding is not simply "help," but a mechanism for preserving learner agency and productive struggle while reducing unproductive confusion.

Engagement, in turn, requires careful conceptual definition. Kahu's (2013) framework in higher education distinguishes engagement as a state arising from interactions among student factors and institutional context, with proximal consequences (e.g., persistence, satisfaction) and distal consequences (e.g., achievement). As a design heuristic, the ICAP (Passive, Active, Constructive, Interactive) framework differentiates Passive, Active, Constructive, and Interactive modes of cognitive engagement, predicting deeper learning as activity shifts toward Constructive and Interactive forms (Chi & Wylie, 2014). For AI tutoring, ICAP implies that "engaging chat" is insufficient; systems should repeatedly elicit learner-generated explanations, intermediate steps, and reciprocal dialogue rather than only delivering polished responses.

Because this paper emphasizes early engagement outcomes, it also draws on HCI (Human–Computer Interaction) measurement work. O'Brien and Toms (2008) define user engagement as a quality of user experience characterized by attention, affect, and endurability, among other attributes. O'Brien et al. (2018) refine engagement measurement through the User Engagement Scale (UES) and short form, providing constructs (e.g., focused attention; perceived usability; endurability) that can inform evaluation design for embedded tutoring interfaces. These constructs are particularly relevant when an AI tutor is integrated into an LMS workflow and adoption depends on both perceived usefulness and a low-friction experience.

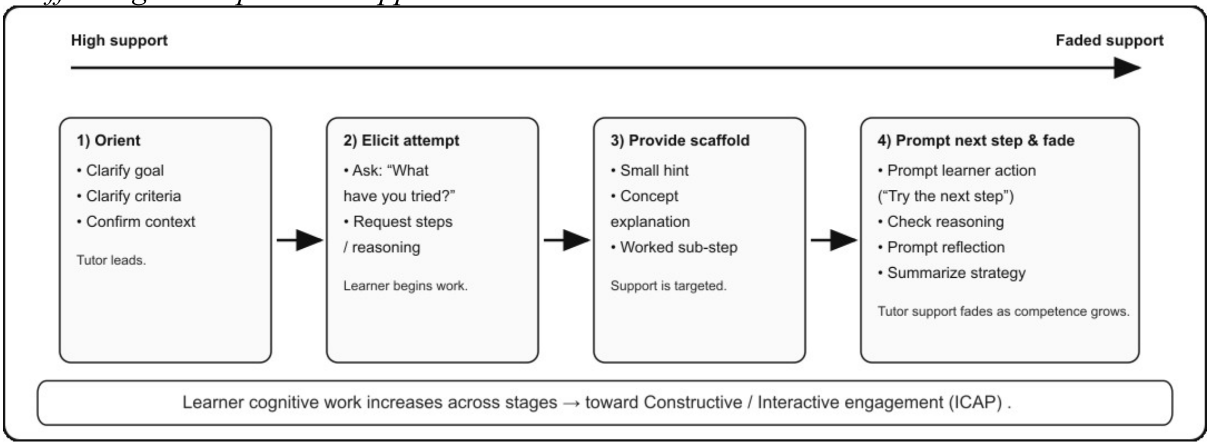
Finally, the development model matters for educational innovation. The Students-as-Partners literature conceptualizes learners and staff as collaborators in educational design, emphasizing reciprocity and inclusive partnership practices; systematic review evidence suggests partnerships are frequently small-scale, undergraduate-focused, and oriented toward teaching and learning enhancement (Mercer-Mapstone et al., 2017). This project adopts that orientation by positioning undergraduate developers as co-creators of an authentic learning support system guided by faculty mentorship.

System Design and Implementation

The AI Tutor was implemented as an LMS-embedded tutoring interface within the Canvas LMS intended to reduce help-seeking friction and to keep students within the learning context where questions arise (e.g., assignment pages and course announcements). The system's instructional goal was not to provide finalized solutions, but to guide students through a scaffolded interaction: clarifying the task, eliciting the student's attempt, offering hints or concept explanations, and prompting the student to complete the next step. This design choice directly targets the "crutch" risk identified in recent evidence: learners may practice less productively when AI provides direct answers, whereas tutorial hinting can preserve skill learning (Bastani et al., 2025).

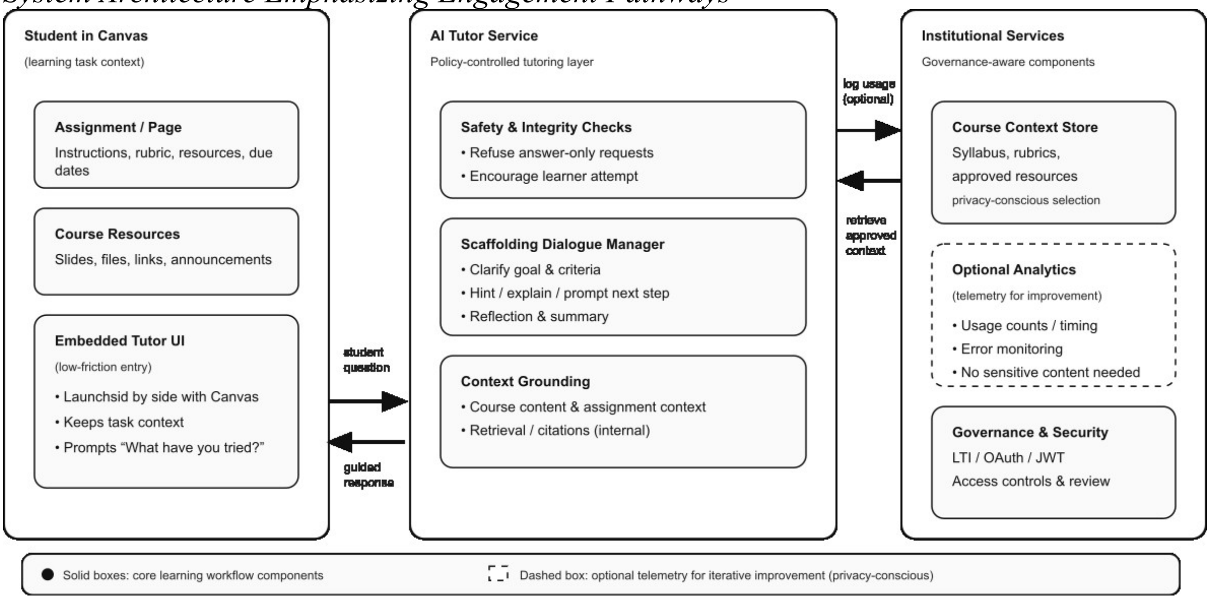
Canvas integration choices also shape adoption and governance. In general, LMS-integrated tools frequently rely on the Learning Tools Interoperability (LTI) standard; Canvas documentation describes "external tools" as IMS LTI links and provides configuration and API mechanisms for deploying tools into courses and accounts (Instructure, n.d.). For modern deployments, LTI 1.3 and LTI Advantage leverage the 1EdTech Security Framework. 1EdTech explicitly describes LTI 1.3 as adopting OAuth 2.0 for authentication services, JSON Web Tokens (JWT) for message signing, and the OpenID Connect workflow paradigm, supporting security and "mobile-ready" interoperability (1EdTech Consortium, n.d.). These standards contextualize institutional choices about safe integration pathways for AI tools operating in learning environments. Figure 1 illustrates the scaffolding conceptual model implemented in the AI Tutor.

Figure 1
Scaffolding Concept Model Applied in the AI Tutor



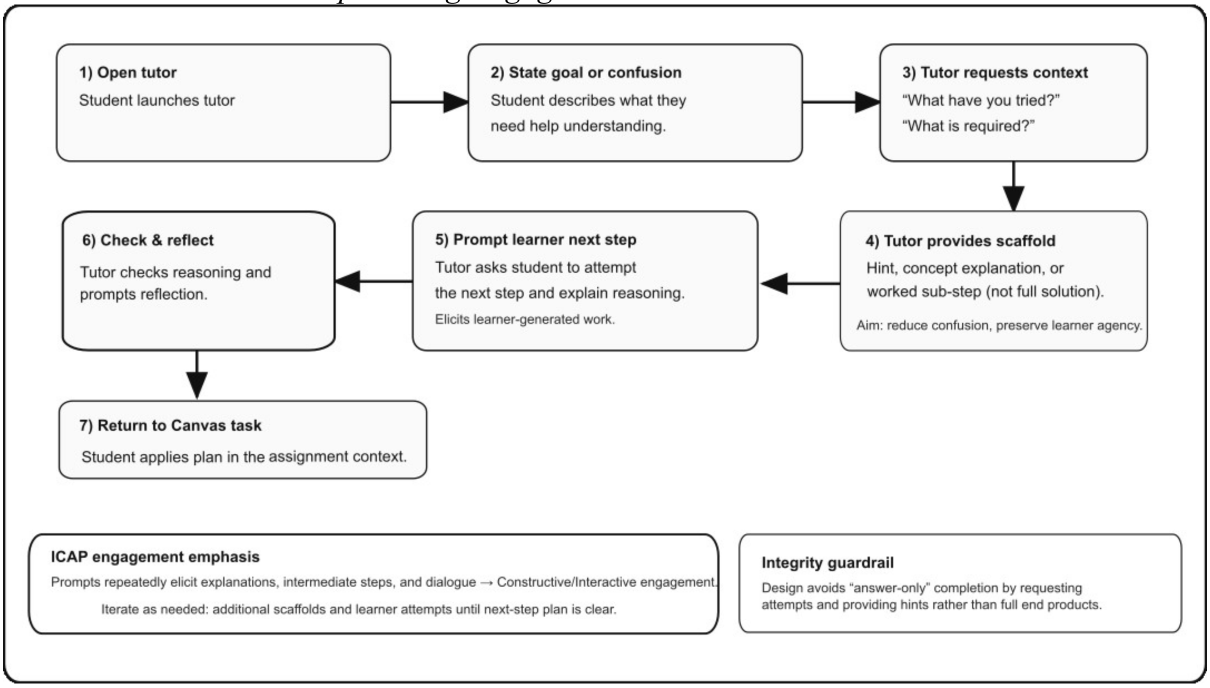
Learner behaviors shift toward Constructive/Interactive activity (ICAP), emphasizing that the learner—not the AI—performs the key cognitive work (Chi & Wylie, 2014; Wood et al., 1976). Figure 2 illustrates the system architecture and engagement pathways.

Figure 2
System Architecture Emphasizing Engagement Pathways



This architecture is designed to support engagement by reducing platform switching and by enabling course-contextual relevance, which prior LMS-embedded assistant research identifies as a key driver of satisfaction and engagement when AI assistance is grounded in course content (Mezwri & Turcsányi-Szabó, 2025). Figure 3 illustrates user interaction flow and engagement.

Figure 3
User Interaction Flow Emphasizing Engagement



This flow is explicitly aligned to foster Constructive/Interactive engagement (ICAP) by repeatedly eliciting learner-generated reasoning rather than delivering end products (Chi & Wylie, 2014).

The AI Tutor was developed by a four-member undergraduate team operating under a faculty-mentored, students-as-partners model. Students led the technical implementation, including backend development, interface design, AI integration, and system testing, while the faculty mentor provided pedagogical framing, project governance guidance, and institutional alignment. Development occurred iteratively over approximately one academic year using agile methods, with students responsible for feature prioritization, troubleshooting, and deployment decisions. Beyond producing a functional tutoring system, the project served as an experiential learning environment in which students developed competencies in AI system integration, full-stack development, and interdisciplinary collaboration.

Methods and Data Sources

This exploratory pilot examined early engagement outcomes during the system's initial classroom deployment. The evaluation used mixed evidence sources appropriate for a preliminary engagement study: (a) post-use perception surveys; (b) system usage indicators (as descriptive evidence of adoption); and (c) student reflections capturing how the tutor fit into help-seeking and problem-solving processes.

This pilot was conducted in introductory level course at Utah Valley University during Fall 2025. Of N = 22 enrolled students, 17 used the tutor at least once. We collected (a) an end-of-pilot survey, (b) system logs (sessions, messages, timestamps).

Survey constructs were aligned with established engagement measurement perspectives (e.g., attention, perceived usability, and endurability), drawing conceptually from the UES tradition rather than claiming full instrument equivalence (O'Brien et al., 2018).

The analysis focus was intentionally limited to early engagement and instructional fit. The study did not treat engagement perceptions as direct proxies for learning gains, and it did not claim causal inferences due to the absence of experimental controls in the pilot setting. This methodological stance is consistent with engagement theory that distinguishes engagement as a proximal state influenced by context rather than a guaranteed indicator of achievement (Kahu, 2013).

Engagement Findings and Interpretation

Descriptive usage data indicated sustained student interaction with the tutor during the pilot. Average session duration was approximately nine minutes, suggesting multi-turn, inquiry-based dialogue rather than brief transactional use. Survey responses further indicated high perceived usability and engagement, consistent with the system's scaffolding-oriented design. These early indicators support the claim that LMS-embedded AI tutoring can achieve meaningful student adoption when aligned with course context.

Across surveys and reflections, students reported high engagement with the system and described it as helpful for understanding course expectations and content. Students emphasized that guided questioning and stepwise support encouraged "thinking through" problems rather than simply retrieving answers. Interpreted through ICAP, these reports suggest movement toward Constructive/Interactive activity—students generating explanations and responding to iterative prompts—rather than Passive consumption.

These early engagement indicators align with broader findings that educational impact is role-dependent: generative AI tools can support learning and learning perceptions when positioned as structured tutoring aids.

Conclusion, Implications, and Future Directions

This paper presented an engagement-centered account of a student-led AI Tutor integrated into Canvas and designed to operationalize scaffolding as a core tutoring behavior. The project's central pedagogical claim is that engagement benefits depend on design: AI tutoring can increase persistence and usability when embedded in context, but it must be constrained to preserve cognitive work and educational integrity. Recent evidence reinforces that unguarded AI can harm learning outcomes after access is removed, while safeguards that support hinting and guided reasoning mitigate this risk (Bastani et al., 2025).

A secondary contribution of this work is demonstrating that student-led AI development can function as a scalable pedagogical model, not only a technology solution.

Implications for higher education include the need to treat AI tutoring as an instructional design decision rather than a generic technology add-on. Engagement should be evaluated for quality, not only quantity. Institutions also need to pair AI tutoring deployments with integrity guidance and assessment redesign, recognizing evidence that detection and authenticity alone may be insufficient safeguards.

Future work should (a) measure engagement quality more directly (e.g., evidence of learner-generated reasoning in dialogues), (b) test scaffold configurations that explicitly push Constructive/Interactive behaviors, (c) evaluate longer-term learning and transfer under different guardrail designs, and (d) formalize governance playbooks aligned with institutional privacy and risk frameworks.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the author used Gemini for proofreading and to improve the writing quality. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the final content of the publication.

References

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, *122*(26), e2422633122. https://doi.org/10.1073/pnas.2422633122

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, *61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Instructure. (n.d.). *External tools API*. Canvas Developer Documentation. https://developerdocs.instructure.com/services/canvas/resources/external_tools

Kahu, E. R. (2013). Framing student engagement in higher education. *Studies in Higher Education*, *38*(5), 758–773. https://doi.org/10.1080/03075079.2011.598505

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., . . . Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kestin, G., Miller, K., Klales, A., Milbourne, T., & Ponti, G. (2025). AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. *Scientific Reports*, *15*, 17458. https://doi.org/10.1038/s41598-025-97652-6

Kofinas, A. K., Tsay, C. H.-H., & Pike, D. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology*, *56*(6), 2522–2549. https://doi.org/10.1111/bjet.13585

Mezwari, K., & Turcsányi-Szabó, M. (2025). Bridging LMS and generative AI: Dynamic course content integration (DCCI) for enhancing student satisfaction and engagement via the ask ME assistant. *Journal of Computers in Education*. Advance online publication. https://doi.org/10.1007/s40692-025-00367-w

Mercer-Mapstone, L., Dvorakova, S. L., Matthews, K. E., Abbot, S., Cheng, B., Felten, P., Knorr, K., Marquis, E., Shammas, R., & Swaim, K. (2017). A systematic literature review of students as partners in higher education. *International Journal for Students as Partners*, *1*(1), 15–37. https://doi.org/10.15173/ijsap.v1i1.3119

O'Brien, H. L., Cairns, P., & Hall, M. (2018). A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human–Computer Studies*, *112*, 28–39. https://doi.org/10.1016/j.ijhcs.2018.01.004

O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, *59*(6), 938–955. https://doi.org/10.1002/asi.20801

1EdTech Consortium. (n.d.). *Why platforms and tools should adopt LTI 1.3*. https://www.1edtech.org/standards/lti/why-adopt-lti-1p3

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., Kyza, E., Edelson, D., & Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences*, *13*(3), 337–386. https://doi.org/10.1207/s15327809jls1303_4

Tabassi, E. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)* (NIST AI 100-1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: Insights from a meta-analysis. *Humanities & Social Sciences Communications*, *12*, 621. https://doi.org/10.1057/s41599-025-04787-y

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Contact email: ahmed.alsharif@uvu.edu