Designing for Student Engagement with AI in Courseware: Lessons from Iterative Improvements to DOT in REAL CHEM

Abigail Stein, Shang-Ting Ciou, Tingyue Cui, Kimberly Larson, Lilly Lee, Sherry Li, Chris Mead, Ashley Xu, & David J. Yaron

AI Courseware learning engineering Student Engagement

This design-based research study reports on a Learning Engineering (LE) cycle embedded within the broader REAL CHEM courseware initiative, examining how students use and perceive DOT, a generative AI tutor for general chemistry. In Fall 2024, student awareness and engagement with DOT were low, and interviews revealed reliance on external AI tools and limited generative use. Guided by LE principles—human-centered design, iterative refinement, and evidence-informed decision making—we redesigned DOT by refining its base prompt and introducing AI Activation Points to make the system more proactive and contextually aligned with student workflows. Summer 2025 findings showed increased awareness, higher engagement, and improved student satisfaction. Results illustrate how targeted, iterative LE cycles can improve the effectiveness of AI-enabled learning supports at scale.

Introduction

REAL CHEM is a fully instrumented courseware environment for general chemistry developed at Carnegie Mellon University and Arizona State University (CMU, 2022) and built around the OpenStax Chemistry 2e textbook (Flowers et al., 2009). REAL CHEM integrates content, assessment, and analytics into a coherent instructional system designed to support student learning, particularly for students historically underserved in STEM.

The design and evolution of REAL CHEM reflect core commitments of Learning Engineering (LE), an interdisciplinary field that integrates learning sciences, instructional design, data, and human-centered methods to iteratively improve learning systems through evidence (Baker et al., 2022; Goodell et al., 2023). Within this larger effort, this paper focuses on a bounded LE cycle centered on DOT, the Digital Online Tutor.

DOT is a generative AI tutor embedded within the REAL CHEM interface. At the outset of this study, DOT functioned as an omnipresent chat tool available to students but rarely used. This paper documents how learner data and interviews were used to diagnose challenges related to awareness, trust, and use; inform targeted design decisions; and evaluate the impact of those decisions.

Methods

Our methods follow a design-based research approach (Cobb et al., 2003) aligned with LE's iterative design-test-refine cycles. In Fall 2024, we conducted semi-structured interviews with 10 undergraduate students enrolled in REAL CHEM-supported

courses and analyzed DOT interaction logs from 473 students. Interview questions focused on awareness of DOT, use of AI tools (e.g., ChatGPT), trust and accuracy perceptions, and decisions about when to seek AI support.

Findings from this cycle informed targeted design updates to DOT. In Summer 2025, we conducted a second round of interviews (n=6) and analyzed interaction logs (n=172 students). Interview protocols were extended to probe student experiences with the redesigned DOT and AI Activation Points.

Student-DOT interactions were coded into categories (e.g., content, course administration, platform issues, feedback, and unserious conversation). Descriptive statistics were used to examine engagement patterns across cycles.

Results

In Fall 2024, student awareness and use of DOT was extremely limited. Half of interviewed students were unaware of the feature, and only 11% of the 4,249 enrolled students interacted with DOT, most fewer than five times. Students reported relying on external AI tools (e.g., ChatGPT, Gemini, Claude) and traditional resources such as peers, instructors, and YouTube. Interaction data showed that queries were dominated by copied course questions (32%) and factual recall (29%), with limited generative or explanatory use. Students also cited concerns about accuracy, verbosity, and poorly rendered mathematical notation.

Based on these findings, we made two design changes. First, we refined DOT's base prompt to prioritize concise, stepwise explanations, avoid directly providing final answers, and reduce mathematical and formatting errors. Second, we introduced AI Activation Points—structured moments where DOT proactively engages students at key points in the courseware—to improve visibility and alignment with student workflows.

Summer 2025 findings showed substantial improvement. All interviewed students were aware of DOT, overall interaction rates increased to 36.2%, and student satisfaction ratings averaged 4.5/5. Activity-level and open-ended Activation Points were perceived as timely and helpful, while page-level activations showed low engagement, suggesting the importance of precise timing. Copying and pasting course questions into DOT decreased.

Discussion

These findings demonstrate how a focused LE cycle can improve student engagement with AI tutors embedded in courseware. Rather than treating DOT as a static feature, learner data and qualitative insights were used to guide targeted, human-centered design decisions and evaluate their impact. Results help explain why students often prefer general-purpose AI tools over course-specific tutors and show how intentional design—proactivity, contextual relevance, and workflow alignment—can narrow that gap.

More broadly, this study illustrates the value of nested LE cycles within large-scale courseware initiatives, enabling systematic improvement of AI-enabled supports while remaining responsive to student needs.

Acknowledgments

We gratefully acknowledge the financial support from The Bill & Melinda Gates Foundation to this research effort. Opinions and conclusions expressed in this article do not necessarily reflect the views of this funding agency.

References

Baker, R. S., Boser, U., & Snow, E. L. (2022). Learning engineering: A view on where the field is at, where it's going, and the research needed. Technology, Mind, and Behavior. https://doi.org/10.1037/tmb0000058

Carnegie Mellon University. (2022, June 8). Universities partner to make chemistry more equitable. https://www.cmu.edu/news/stories/archives/2022/june/asu-cmu-chemistry.html

Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. Educational researcher, 32(1), 9-13.

Flowers, P., Theopold, K., Langley, R., & Robinson, W. (2019). Chemistry 2e. OpenStax. https://openstax.org/details/books/chemistry-2e

Goodell, J., Kessler, A., & Schatz, S. (2023). Learning Engineering at a Glance. Journal of Military Learning. https://www.armyupress.army.mil/Portals/7/journal-of-military-learning/images/Conference-Edition-2023/JML-Conference-Edition-2023-TOC-v1.pdf

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Guggemos, J., Opfermann, M., Schmid, S., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Instruction, 88, 101666.