

## Training AI Tutors: Tutor Training of the 21st Century

Eva Politou  
Harvard University

Megan Selbach-Allen  
Harvard University

Ari Kaufman  
Harvard University

Milan Naropanth  
Harvard University

*Keywords:* Tutor Training, Multivariable Calculus, AI Tutor

### Introduction and Theoretical Framework

AI tools like ChatGPT and other Large Language Models (LLM) are here to stay, and already in use in educational platforms like Khanmigo. These tools show a lot of potential for use in education, but in their current form appear to be most useful for semi-supervised content generation (Nye et al., 2023). While AI tutors have potential to help students learn, provide personalized guidance and expand educational support they still require training and supervision. Using a lens of undergraduate tutor training (Johns, 2024) we sought to answer the questions: How does an AI tutor respond to training and how effectively does it support students?

### Methods

This study was a case study of an AI tutor developed to support students in a multivariable calculus course. Students had the option of attending an in-person problem solving session led by an undergraduate course assistant or engaging with the same problems assisted by an AI tutor. The tutor was built on the TeachGPT platform which integrates ChatGPT 4 and 4o into educational tools for personalized learning (Kestin et al., 2024). The AI tutor was trained to be a friendly and encouraging persona focused on inquiry-based learning, guiding students through real-life applications of multivariable calculus. It was directed to avoid giving direct answers, instead offering hints or guiding questions like, "What do you think would happen if...?" to foster critical thinking and reflection. For each workshop, problems and solutions with detailed instructions, including learning objectives, solution steps, common misconceptions, and strategies to address them, creating a supportive and engaging learning environment were uploaded to the platform. Data for this study were the training prompts and logs of student tutor interactions over the course of the semester.

### Results

After initial training, the AI tutor was helpful to students and did not appear to hallucinate or provide incorrect information. However, it appeared to be "too helpful" in that it was likely to provide a complete solution to the problem after a short question from the student. Further training from the course instructor was needed to direct the AI tutor to provide shorter hints and guide students via questioning rather than providing full solutions initially.

### Discussion and Implications

This study demonstrates one potential use of LLMs in an undergraduate mathematics classroom. The course instructor felt the fully trained model performed well as a tutor and this modality was popular with students although many students still preferred to attend in-person sessions. Further research is needed to provide more general guidance for effective training of LLMs to act as mathematics tutors for students. With the current state of this technology the role of the human instructor remains central to the success of the AI tutor.

### References

- Nye, B. D., Mee, D., & Core, M. G. (2023). Generative Large Language Models for Dialog-Based Tutoring: An Early Consideration of Opportunities and Concerns. In *AIED Workshops* (pp. 78-88).
- Johns, C. (2024). Research-based training for undergraduate mathematics tutors. *International Journal of Mathematical Education in Science and Technology*, 55(10), 2410-2435.
- Kestin, G., Miller, K., Klales, A., Milbourne, T., & Ponti, G. (2024). AI Tutoring Outperforms Active Learning. Research Square. <https://doi.org/10.21203/rs.3.rs-4243877/v1>