Generative AI and the Protégé Effect in the Classroom

Jonathan Steffens and Charles Pezeshki

Washington State University, Department of Mechanical and Materials Engineering

Abstract

This paper presents the pilot implementation of a Custom GPT in ME304 - Heat Transfer, at Washington State University. The custom GPT harnesses the "learning by teaching" pedagogical method, also known as the protégé effect. By instructing students to regularly teach the GPT course topics while it only asked clarifying follow-up questions, this approach helped learners actively reflect on their reasoning, reinforce core concepts, and uncover misconceptions. Preliminary survey results indicate that students found the interactions both engaging and beneficial for conceptual mastery, suggesting that Custom GPTs can offer an effective and easy to implement means to support complex STEM courses. This proof-of-concept lays the groundwork for broader adoption of AI-driven pedagogical strategies in STEM education, with potential for refining domain-specific prompts and integrating other adaptive learning tools in future implementations.

1. Introduction

Teaching a subject to others has long been recognized as an effective strategy for reinforcing one's own understanding. This is sometimes referred to as the protégé effect. This method helps to identify misconceptions, deepen engagement with the material, and ultimately achieve a more robust comprehension [1] [2] [3]. Traditionally, this effect has been studied in peer-to-peer environments, especially in tutoring programs [4] [5]. Early advances in AI agents and other digital tools enabled some previous effective uses of this technique [6] [7] [8], often in the form of rule-based teachable agents that prompted students to explain their reasoning.

However, recent advances in large language models (LLMs) present new, low-cost, and easy-to-implement opportunities to deploy this method in classroom settings at scale. These tools are capable of simulating natural, dynamic conversation, allowing students to interact with an AI agent that feels more like a genuine novice learner than previous static systems. These capabilities build on a growing body of work demonstrating that teachable agents — particularly those that can engage in dialogue — can foster metacognition, promote active engagement, and lead to measurable learning gains [3].

This is especially relevant for STEM education, where complex problem-solving and conceptual reasoning are often challenging for students. Recent studies have shown that positioning students as teachers to an LLM-based agent can improve not only content knowledge, but also critical higher-order skills such as reflection, self-regulation, and monitoring of their own understanding [9]. By requiring students to organize their knowledge, explain it clearly, and respond to probing questions, these systems encourage a more thoughtful and structured engagement with the material.

Moreover, generative AI tools offer distinct advantages over human-to-human teaching opportunities. Peer tutoring, while highly effective, is logistically difficult to implement broadly in large university courses. Faculty and teaching assistants have limited time for one-on-one dialogue with every student. AI-driven teachable agents, by contrast, are scalable, always available, and capable of individualized interaction for every learner [10]. This flexibility has been shown to support greater student autonomy, self-paced exploration, and increased confidence when working through difficult material [11].

In the context of STEM fields, where students often struggle with conceptual clarity or become overwhelmed by procedural problem-solving, these benefits are particularly impactful. Studies using LLM-based AI in subjects like programming [3], and mathematics [11] have demonstrated improvements in both test performance and the integration of knowledge. Students are not simply performing isolated steps but are developing more cohesive mental models that connect core principles.

Language models and AI agents are thus an effective partner for students to "teach" because they encourage consistent, reflective explanations of ideas, prompting learners to refine and clarify their understanding in the process. Perhaps most importantly, these agents offer individualized, around-the-clock access for students in a way faculty, teaching assistants, and tutors often cannot provide.

ChatGPT from OpenAI [10] offers a way to create customized language models which can be tailored to effectively implement behavior to mimic a learner for students to interact with Custom GPTs. These use retrieval augmented generation (RAG) to provide external knowledge sources to supply accurate, domain-specific information and enhance the realism of the learner–student interactions [12]. The creation of a Custom GPT requires two things: 1. Domain-specific knowledge, such as notes and example problems, and 2. Instructions for how to respond to and interact with users.

Given these emerging possibilities, the present work explores a pilot implementation of a Custom GPT in ME304 – Heat Transfer – at Washington State University. This study investigates whether integrating a GPT-based learning bot, designed to invoke the protégé effect, can provide a simple, scalable method to enhance conceptual engagement in a complex engineering course.

2. Methodology

A pilot study for the use of Custom GPTs to enact the protégé effect has been implemented in ME304, Heat Transfer, at Washington State University. A custom GPT called "learning bot" was given full instructor notes from the course as well as worked example problems and Python code used to solve various kinds of problems. In addition, the following instructions were given:

"You are a curious undergraduate student in mechanical engineering. You are intelligent but you don't know very much, and you want to learn! Ask the user to teach you something. Ask relevant and thought provoking follow up questions. Ask for clarity when the explanations are not well written or conceived.

Do not supply new information. If something is unknown, ask the student to find out and then come back later. It's not your job to explain anything. After four rounds of conversation, thank the user for teaching you something new and interesting.

You should be especially interested and ask questions about temperature, heat, and heat transfer mechanisms like conduction, convection, and radiation as well as practical applications to mechanical engineering."

Students were required to interact with the learning bot on a weekly basis. They were asked to teach it something related to the week's topics. They were encouraged to make the topic more specifically related to something of personal interest. The hope is that they can better connect the lecture material to subjects which they already have some knowledge of to better solidify the applicability of heat transfer topics to myriad situations.

3. Results and Discussion

This approach has, thus far, only been implemented in a single course over one semester. Consequently, it is not intended to provide a comprehensive analysis of efficacy but rather to serve as a preliminary proof-of-concept for the implementation of Custom GPT technology in the classroom.

A student survey was conducted using two Likert-scale prompts and one open-ended question, asking participants to rate their experience with the learning bot. Of the 50 students in the class, 36 responded. The specific survey prompts are as follows:

Q1: “The time I spent interacting with the learning bot was valuable to my overall learning process.”

Q2: “The bot’s follow-up questions were insightful and relevant to the conversation.”

Q3: "Provide any additional comments about your experience with the learning bot."

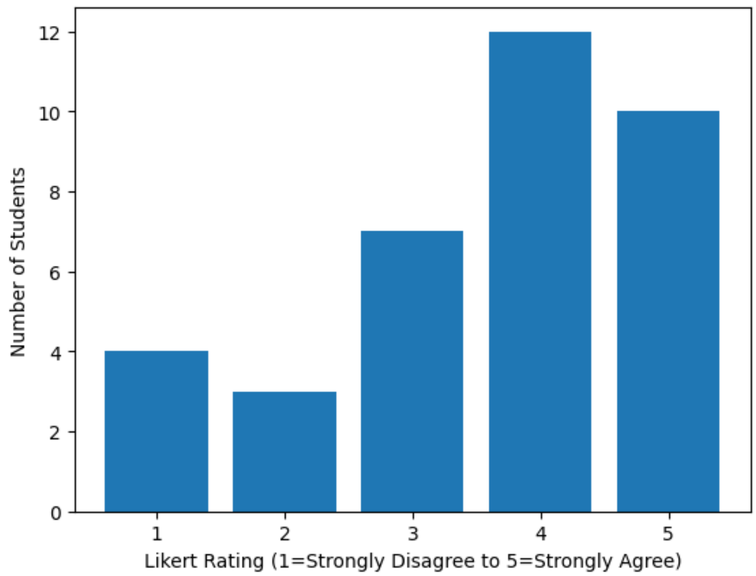

Figure 1: Responses to Q1 - “The conversation with the learning bot helped me clarify key course concepts.”

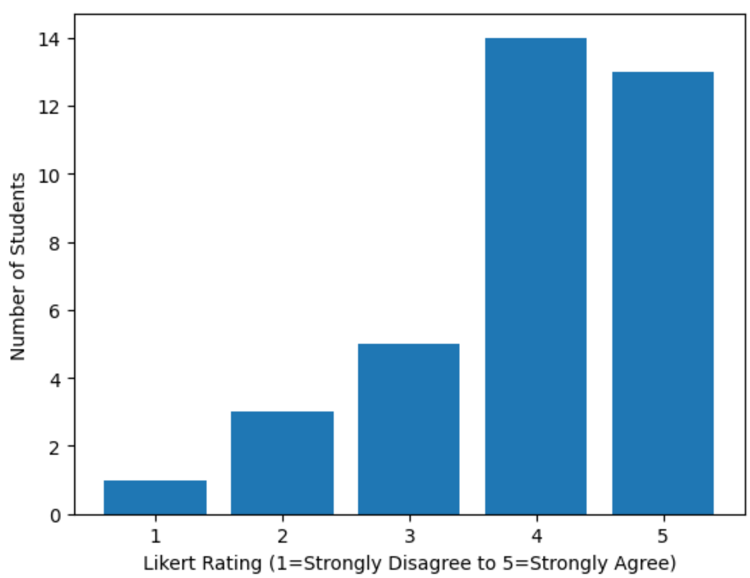

Figure 2: Responses to Q2 - “The bot’s follow-up questions were insightful and relevant to the conversation.”

Most students responded positively to their interactions with the learning bot and found it helpful in clarifying key concepts. Students also found the learning bot to ask relevant and engaging follow up questions. General comments were positive overall and some even expressed surprise about how useful they found these interactions.

There are several negative responses, though it is unclear the major drivers of this dissatisfaction. One possible issue is a general dislike of AI tools by some students.

4. Conclusion

This pilot implementation of a Custom GPT in ME304, Heat Transfer, demonstrates the potential benefits of leveraging the protégé effect through virtual interactions. By requiring students to serve as "teachers" to the learning bot, the approach promotes deeper engagement with course content and encourages consistent reflection, review, and articulation of key concepts. Though limited to a single semester, survey responses indicate that the guided questioning strategy supported students' comprehension and self-awareness of their own knowledge gaps. Moreover, the accessible, scalable nature of Custom GPTs offers promise for seamless integration into diverse educational contexts, extending the reach of student-centered, active learning approaches. Future research will focus on scaling this method across multiple courses, refining the design of domain-specific prompts, and investigating the long-term impact on academic performance and student motivation.

5. Future Work

While the pilot study presented here demonstrates the potential benefits of utilizing Custom GPT technology to harness the protégé effect, a more robust measurement approach is necessary to assess long-term impacts on student performance and conceptual understanding. Furthermore, there is significant opportunity to integrate additional AI-driven pedagogical methods, such as adaptive assessments and automated feedback systems, to more thoroughly engage and challenge students. By combining these tools, instructors can create a more individualized, interactive learning environment that supports different learning styles and paces. Ultimately, continued exploration in this direction will help establish best practices, demonstrate the efficacy of AI-facilitated teaching strategies, and pave the way for more seamless, widespread adoption in STEM education.

References

[1] Roscoe, R. D., & Chi, M. T. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. Review of Educational Research, 77(4), 534–574.

[2] Cohen, P. A., Kulik, J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. American Educational Research Journal, 19(2), 237–248.

[3] Chen, A., Wei, Y., Le, H., & Zhang, Y. (2024). Learning-by-teaching with ChatGPT: The effect of teachable ChatGPT agent on programming education. *arXiv preprint arXiv:2412.15226*.

[4] Topping, K. J. (2005). Trends in peer learning. Educational Psychology, 25(6), 631–645.

[5] Goodlad, S., & Hirst, B. (1989). Peer Tutoring: A guide to learning by teaching. London: Kogan Page.

[6] Biswas, G., Leelawong, K., Schwartz, D., & Vye, N. (2005). Learning by teaching: A new agent paradigm for educational software. Applied Artificial Intelligence, 19(3–4), 363–392.

[7] Chase, C. C., Chin, D. B., Oppezzo, M., & Schwartz, D. L. (2009). Teachable agents and the protégé effect: Increasing cognitive gains by teaching agents. Journal of Science Education and Technology, 18(4), 334–352.

[8] Schwartz, D. L., Blair, K. P., Biswas, G., Leelawong, K., & Davis, J. (2007). Animations of thought: Interactivity in the teachable agent paradigm. In R. Lowe & W. Schnotz (Eds.), Learning with animation: Research implications for design. Cambridge University Press.

[9] Chang, C. C., Chang, C. K., & Shih, J. L. (2016). Motivational strategies in a mobile inquiry-based language learning setting. System, 59, 100–115.

[10] OpenAI. (2023). GPT-4 Technical Report. Accessed at: https://cdn.openai.com/papers/GPT-4.pdf

[11] Pardos, Z. A., & Bhandari, S. (2023). Learning gain differences between ChatGPT and human tutor generated algebra hints. *arXiv preprint arXiv:2302.06871*.

[12] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-T., & Riedel, S. (2020). Retrieval-augmented generation for knowledge-intensive NLP. In Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS). arXiv:2005.11401