Practical evaluation of the possibilities of integrating Large Language Models in physics laboratory instruction

Marina BABAYEVA (1), Travis KREGEAR (2), Ralf WIDENHORN (3)

(1) Charles University, Faculty of Mathematics and Physics, Department of Physics Education, V Holešovičkách 2, 180 00 Praha 8, Prague, Czech Republic
(2) Portland State University, Department of Physics, 1719 SW 10th Ave., Portland, OR 97201
(3) Portland State University, Department of Physics, 1719 SW 10th Ave., Portland, OR 97201

Abstract. A team of researchers from Portland State University (PSU) and Prague's Charles University pioneered the integration of OpenAI's GPT-4 Large Language Module (LLM) into the introductory physics laboratory, serving as a virtual teaching assistant. A class of 26 students participated in a pilot implementation of the AI assistant for introductory physics at PSU. We evaluated students' interactions with the chatbot, assessed the quality of the LLM's responses, and conducted surveys to gather qualitative insights from students' experiences and attitudes. This contribution presents the key findings of our pilot study, including examples from the transcripts and surveys.

Current state

Since the launch of OpenAI's ChatGPT [1], Large Language Models (LLMs) have received significant public attention due to their ability to generate coherent and contextually relevant responses on just about every topic. In a relatively short time, various AI tools have gained widespread use. Thus, the research of AI applications in the field of physics education is accelerating in multiple directions. Physics education researchers have studied student perception of ChatGPT [2], reviewed trends of the integration of LLM in education [3], compared results for prospective physics teachers [4], and developed strategies [5,6] and frameworks [7] for incorporating generative AI in education. LLM models were tested to answer multiple physics exam questions on different levels of complexity [8,9]. For STEM and non-STEM disciplines, researchers have shown the use of generative AI can increase critical thinking and problem-solving, both skills that are important to gaining physics excellence [10,11]. What is clear at this point, generative AI provides opportunities, risks, and challenges for physics education [12].

Pilot study

A team of researchers from Portland State University (PSU) and Charles University developed the interface of the AI-powered lab assistant. The primary goal of this assistant is to provide immediate feedback and support to students as they work through introductory physics lab activities. This is achieved by combining student responses with instructional prompts and sending them via API call to the LLM GPT-4 model, and then receiving AI feedback. The assistant was initially implemented for laboratory work in an introductory physics laboratory course for a lab on the moment of inertia. A total of 26 students participated in the pilot implementation at the Department of Physics at PSU. The study aims to answer the following questions:

- What are the challenges of implementing LLMs in an introductory physics lab setting?
- From a student perspective, how helpful was the LLM support?
- How accurate and pedagogically effective was the information provided by the LLM?

The research involved analyzing student interactions with the LLM during lab sessions. The team evaluated the quality of AI feedback and students' attitudes before and after the pilot and collected the following data:

- Students' pre- and post-surveys.
- Students - AI interaction transcript.

Anticipated risks included the possibility of vague and misleading answers that could lead to incorrect conclusions or judgments. Collected data was analyzed and assessed to establish a performance baseline for the LLMs using statistical models (including sentiment and engagement). The outcomes of the preliminary project offered us valuable information on the present capabilities of AI in physics laboratories. It lays the foundation to study the possibilities and challenges of future research advancements in this field.

References

[1] *ChatGPT*. (2024, February 20). OpenAI. https://openai.com/chatgpt

[2] P. Bitzenbauer, ChatGPT in physics education: A pilot study on easy-to-implement activities. *Cont. Educ. Tech.* **15**(3) (2023) ep430. https://doi.org/10.30935/cedtech/13176

[3] A. Boltsi, K. Kalovrektis, A. Xenakis, P. Chatzimisios, C. Chaikalis, (2024). Digital tools, technologies, and learning methodologies for education 4.0 Frameworks: A STEM oriented survey. *IEEE Access* **12** (2024) 12883–12901. https://doi.org/10.1109/access.2024.3355282

[4] S. Küchemann, S. Steinert, N. Revenga, M. Schweinberger, Y. Dinc, K. E. Avila, J. Kuhn, Can ChatGPT support prospective teachers in physics task development? *Physical Rev. Phys. Educ. Res.* **19**(2) (2023) https://doi.org/10.1103/physrevphyseducres.19.020128

[5] K. Hettiarachchilagea, N. Haldolaarachchige, Effective model with personalized online teaching and learning science in the era of chatgpt. arXiv.Org. (2023, June 15). https://arxiv.org/abs/2306.09545

[6] C. Zhu, M. Sun, J. Luo, T. Li, M. Wang, How to harness the potential of ChatGPT in education? *Knowledge Management & E-Learning: An International Journal* **15**(2) (2023) 133–152. https://doi.org/10.34105/j.kmel.2023.15.008

[7] J. Su (苏嘉红) and W. Yang (杨伟鹏), Unlocking the Power of ChatGPT: A framework for applying generative AI in education, *ECNU Review of Education* **6**(3) (2023) 355–366. https://doi.org/10.1177/20965311231168423

[8] W. Yeadon and T. Hardy, *The impact of AI in physics education: A comprehensive review from GCSE to university levels*. arXiv.Org. (2023, September 10). https://arxiv.org/abs/2309.05163

[9] X.-Q. Dao, N.-B. Le, X.-D. Phan, B.-B. Ngo, An evaluation of ChatGPT's proficiency in English language testing of the Vietnamese national high school graduation examination. *SSRN Electronic Journal* (2023) https://doi.org/10.2139/ssrn.4473369

[10] R. Yilmaz and F. G. Karaoglan Yilmaz, The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence* **4** (2023) 100147. https://doi.org/10.1016/j.caeai.2023.100147

[11] Q. Xia, T. K. F. Chiu, C. S. Chai, K. Xie, The mediating effects of needs satisfaction on the relationships between prior knowledge and self-regulated learning through artificial intelligence chatbot. *British Journal of Educational Technology* **54**(4) (2023) 967–986. https://doi.org/10.1111/bjet.13305

[12] Md. M. Rahman and Y. Watanobe, ChatGPT for Education and Research: Opportunities, threats, and strategies, *Appl. Sci.* **13**(9) (2023) 5783. https://doi.org/10.3390/app13095783