Exploring the Use of GPT-4 in Co-creating Personalized Case Scenarios for Higher Education.

Pablo Flores[1*], Guang Rong[1] and Benjamin Ultan Cowley[1,2]

[1*]Faculty of Educational Sciences, University of Helsinki,
Siltavuorenperger 5, Helsinki, 00014 , Finland.
[2*]Cognitive Science, Faculty of Arts, University of Helsinki.

*Corresponding author(s). E-mail(s): pablo.flores@helsinki.fi;
Contributing authors: guang.rong@helsinki.fi ; ben.cowley@helsinki.fi;

Abstract

In the fast evolving landscape of Artificial Intelligence in Education (AIEd) Large Language Models (LLMs) like GPT-4 have emerged as powerful and versatile tools capable of adapting for a wide set of natural language tasks. This study explores into an interaction design where higher education students utilized GPT-4 to craft personalized case scenarios tailored for their coursework. This process not only facilitates tailored learning experiences, but also engaged students into speculative reflections about the potential futures of AIEd ..... Add up on results and conclusions when done

Keywords: ChatGPT, Large Language Models, Higher Education, Human-AI Co-creation ,Personalized Learning, Speculative Scenarios

1 Introduction

Upon every technological breakthrough, the educational domain reconsiders its methods and strategies. During the recent decades, the rise of digital technologies has reshaped the educational landscape. Evidence from the past 40 years on educational technologies impact consistently indicates positive benefits from its integration in programming courses or through innovative approaches (Higgins et al., 2012). However, while digital technologies might spark motivation and engagement in young students, its mere adoption does not ensure favorable outcomes. To fully embrace their potential, a deliberate pedagogical approach is essential, especially when transitioning from

traditional to modern educational approaches (Khaddage et al., 2021; Parker et al., 2020) .

As part of the larger digital transformation, Artificial-Intelligence (AI)-based systems have emerged as important players in the educational domain. These systems have found extensive use in administration, instruction, and learning (Chen et al., 2020), accompanied by a growing body of research around AI in Education (AIEd), a distinct sub-field of digital learning research (Niemi et al., 2023). But despite the wide range of potential applications of AI-based systems, they have been mostly implemented to facilitate, or automate, mainstream learning approaches. While useful, this approach sidelines teachers agency, experience, and creativity to integrate such technologies into their unique practices and pedagogical designs. (Holmes et al., 2023)

A dominant trend has been that AIEd implementations are predominantly designed by technical teams or departments, such as those focused in Science, Technology, Engineering, and Math (STEM), leading to a notable under-representation of educational or psychological perspectives in AIEd research (Holmes & Tuomi, 2022; Zawacki-Richter et al., 2019). Among the probable causes of this we can highlight the lack of technical expertise, among others, which is a big barrier for educational adoption of technologies (Reid, 2014). However, as with other digital technologies, AI systems are evolving to become more accessible by reducing their technical entry barriers, both for casual use and for the design of new applications.

One of such evolution are Large Language Models (LLMs) such as OpenAI's Generative Pre-Trained (GPT) model. They have significantly evolved into systems capable of performing various natural language tasks (Brown et al., 2020). Interestingly, they have shown emergent features as they are able to perform tasks that were not originally part, or expected, of the design (Wei et al., 2022). Their ability to adapt for different tasks without requiring exhaustive re-design of their architectures, combined with their potential to serve as building blocks for task-specific AI-tools, has led to some researchers to classify them as a general purpose technology (Eloundou et al., 2023). Notably, GPT-3 could adapt to new tasks using an approach known as *In-context learning*—relying on natural language descriptions of the task—an ability that it's predecessor, GPT-2 (with fewer training parameters and data) struggles to perform consistently (Bommasani et al., 2022; Wei et al., 2022). Its last iteration, GPT-4, brought significant increase in performance across every measured metric (OpenAI, 2023). These capabilities present wide opportunities in both design and application. Given that state-of-the-art models do not require advanced technical programming skills, professionals from different domains might now tailor customized tools that align closely with their contexts, needs, and specialized approaches (Cain, 2023). Andrew Ng, a leading expert in the field has stated recently that a new breed of prompt-based AI applications might be designed in a few hours, even minutes, just by providing natural language descriptions of the tasks (Stanford Online, 2023) and this study will exemplify such developments. Altogether, a new wave of LLMs-based AI applications is on the horizon, with many speculated to be influential in the near to mid future.(Bommasani et al., 2022; Bubeck et al., 2023).

Due to its versatility and accessibility, OpenAI's "ChatGPT" has become the fastest-growing application in history since its launch on late 2022 (Milmo, 2023). As

a conversational chatbot, it uses natural language processing to understand and generate human-like text in a dialectical fashion. By providing certain input (prompts) ChatGPT can answer different kinds of questions. Additionally, by using advanced prompting methods users can improve its performance in a wide range of tasks (Fernando et al., 2023; Wei et al., 2023). Potential educational uses involve teaching preparation (generation of course materials, providing suggestions); Assessment (generation of exercises or case scenarios, providing feedback); Learning support (answering questions, summarising information) among others (Lo, 2023; Montenegro-Rueda et al., 2023). However, educational research remains sparse and focuses predominantly in theoretically exploring its potential and limitations (Cain, 2023; Qadir, 2022), and assessing its performance on traditional assessment methods (Nisar & Aslam, 2023). Due to its novelty, a gap in exploratory empirical studies is evident.

We focused our study on exploring the interaction between students and GPT-4 for the creation of personalized course materials in a higher education doctoral course. Our research not only aims to pave the way for innovative designs that enrich and personalize in-person learning experiences using AI tools but also to inspire deeper exploration into the vast potentials of integrating AI in educational settings.

1.1 Approach

The study was conducted in an on-campus doctoral course titled "Basics of AI in education", designed to explore and discuss both historical and contemporary developments of AIEd. It covered an overview of the historical technical evolution of AIEd systems, examination of current popular systems, like AI-tutors and GPT models, a review of the intersection between AI and cognitive sciences, as well as a discussion of emerging ethical concerns and regulatory developments.

The course assignments were designed to explore hypothetical scenarios about AI implementations in education. For this purpose the students where asked to individually design a study, and write a reflective essay about a specific chosen scenario. Consequently, the course topics were broad and suggestive, not focused on any predefined scenario, giving the students freedom to choose a scenario of their own interest.

To facilitate this, we designed a guided interaction for the students to co-create personalized hypothetical scenarios through GPT-4, which involved pre-defined prompting templates and sets of conceptual keywords to construct final scenario generation prompts.

This design idea addresses two purposes in our study.

First, since we had a variety of GPT usage experience—and lack of it—it guided students through their scenario co-creation process.

Secondly, allows for methodical reproducible examination of the interactions, since it provides fixed rules on how to prompt. Consequently, we aim to both facilitate an easier use of GPT, balancing the students' prior experiences with it, and to examine the behavioral aspects of the interaction.

Considering the conversational nature of ChatGPT, where users can prompt repeatedly for more information, we guided our design based on two theoretical frameworks, Information Foraging and Computational Thinking.

With the purpose of gathering empirical data in a controlled fashion, and orienting the students through the use of GPT-4, we designed a specific guided interaction that consisted of a set of prompting templates and keywords to be used. By doing this, we aimed to restrict our research focus on the students behavior within a limited set of actions. To conceive the design we had two guiding frameworks, Information Foraging and Computational thinking

Observing the capabilities mentioned above, we found it intriguing to explore a way that the students could craft their hypothetical scenarios with the use of GPT-4.

It also aims to engage educational specialists in speculative reflection about the future of AI in education. Our research not only aims to pave the way for innovative designs that enrich and personalize physical learning experiences using AI tools but also to inspire deeper exploration into the vast potentials of integrating AI in educational settings.

In this study, we explore the gathered behavioral data from the participants' interaction with GPT-4, it's associations with their self-reported Computational Thinking Skills, and their speculations about potential futures in AIEd.

RQ1. How does the students' Computational Thinking Skills influence their co-creation behavioral interactions with GPT-4

Alternatives:

(i) Through which behaviors do students showcase their Computational Thinking skills during their engagement with GPT-4?

(ii) How are students' Computational Thinking skills reflected in their interactions with GPT-4?

(iii) In what ways do students demonstrate Computational Thinking skills when engaging with GPT-4 in a structured setting?

RQ2. What kind of educational research do educational specialists devise for the application of Artificial Intelligence in education?

Alternatives

(i) What are the predominant themes or focuses in the research envisioned by educational specialists regarding Artificial Intelligence in education?

RQ3. What kind of pedagogical value, or lack of it, do educational specialist envision in the development and integration of AI in education?

Alternatives

(i) To what extent do educational specialists foresee AI influencing pedagogical outcomes in educational contexts?

(ii) What are the anticipated pedagogical benefits and challenges of AI integration as envisioned by educational specialists?

Note, alternatives chosen by a set of alternatives co-created with GPT-4

2 Methods

2.1 Design

2.1.1 Prompting templates

2.1.2 Keywords

2.1.3 Interaction

2.2 Quantitative analysis - Network Analysis

2.3 Qualitative analysis- Content Analysis

3 Results

4 Discussion

Discussions should be brief and focused. In some disciplines use of Discussion or 'Conclusion' is interchangeable. It is not mandatory to use both. Some journals prefer a section 'Results and Discussion' followed by a section 'Conclusion'. Please refer to Journal-level guidance for any specific requirements.

4.1 Limitations

4.2 Further Lines of Research

4.3 Conclusion

Conclusions may be used to restate your hypothesis or research question, restate your major findings, explain the relevance and the added value of your work, highlight any limitations of your study, describe future directions for research and recommendations.

In some disciplines use of Discussion or 'Conclusion' is interchangeable. It is not mandatory to use both. Please refer to Journal-level guidance for any specific requirements.

References

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., ... Liang, P. (2022, July). On the Opportunities and Risks of Foundation Models [arXiv:2108.07258 [cs]]. https://doi.org/10.48550/arXiv.2108.07258

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020, July). Language Models are Few-Shot Learners [arXiv:2005.14165 [cs]]. https://doi.org/10.48550/arXiv.2005.14165

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T., & Zhang, Y. (2023, March). Sparks of Artificial General Intelligence: Early experiments with GPT-4 [arXiv:2303.12712 [cs]]. Retrieved March 31, 2023, from http://arxiv.org/abs/2303.12712

Cain, W. (2023). GPTeammate: A Design Fiction on the Use of Variants of the GPT Language Model as Cognitive Partners for Active Learning in Higher Education, 1293–1298. Retrieved April 18, 2023, from https://www.learntechlib.org/primary/p/221996/

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, *1*, 100002. https://doi.org/10.1016/j.caeai.2020.100002

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023, March). GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models [arXiv:2303.10130 [cs, econ, q-fin]]. https://doi.org/10.48550/arXiv.2303.10130

Fernando, C., Banarse, D., Michalewski, H., Osindero, S., & Rocktäschel, T. (2023, September). Promptbreeder: Self-Referential Self-Improvement Via Prompt Evolution [arXiv:2309.16797 [cs]]. https://doi.org/10.48550/arXiv.2309.16797

Higgins, S., Xiao, Z., & Katsipataki, M. (2012, November). *The Impact of Digital Technology on Learning: A Summary for the Education Endowment Foundation. Full Report* (tech. rep.) (Publication Title: Education Endowment Foundation ERIC Number: ED612174). Education Endowment Foundation. Retrieved November 1, 2023, from https://eric.ed.gov/?id=ED612174

Holmes, W., Bialik, M., & Fadel, C. (2023, January). Artificial intelligence in education. https://doi.org/10.58863/20.500.12424/4276068

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, *57*(4), 542–570. https://doi.org/10.1111/ejed.12533

Khaddage, F., Norris, C., Soloway, E., & Davidson, A.-L. (2021). Bridging the Gap Between Traditional Pedagogies and Digital Technologies "Challenges and Misalignments", 80–87. Retrieved November 6, 2023, from https://www.learntechlib.org/primary/p/219642/

Lo, C. K. (2023). What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature [Number: 4 Publisher: Multidisciplinary Digital Publishing Institute]. *Education Sciences*, *13*(4), 410. https://doi.org/10.3390/educsci13040410

Milmo, D. (2023). ChatGPT reaches 100 million users two months after launch. *The Guardian*. Retrieved November 6, 2023, from https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app

Montenegro-Rueda, M., Fernández-Cerero, J., Fernández-Batanero, J. M., & López-Meneses, E. (2023). Impact of the Implementation of ChatGPT in Education:

A Systematic Review [Number: 8 Publisher: Multidisciplinary Digital Publishing Institute]. *Computers*, *12*(8), 153. https://doi.org/10.3390/computers12080153

Niemi, H., Pea, R. D., & Lu, Y. (Eds.). (2023). *AI in Learning: Designing the Future*. Springer International Publishing. https://doi.org/10.1007/978-3-031-09687-7

Nisar, S., & Aslam, M. S. (2023, January). Is ChatGPT a Good Tool for T&CM Students in Studying Pharmacology? https://doi.org/10.2139/ssrn.4324310

OpenAI. (2023). Gpt-4 technical report.

Parker, J., Maor, D., & Herrington, J. (2020). Authentic online learning: Aligning learner needs, pedagogy and technology [Publisher: Institutes for Educational Research in NSW, SA and WA]. *Issues in Educational Research*, *23*(2), 227–241. https://doi.org/10.3316/ielapa.354613892620670

Qadir, J. (2022, December). Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education. https://doi.org/10.36227/techrxiv.21789434.v1

Reid, P. (2014). Categories for barriers to adoption of instructional technologies. *Education and Information Technologies*, *19*(2), 383–407. https://doi.org/10.1007/s10639-012-9222-z

Stanford Online. (2023, August 30). *Andrew ng: Opportunities in ai - 2023*. https://www.youtube.com/watch?v=5p248yoa3oE

Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metzler, D., Chi, E. H., Hashimoto, T., Vinyals, O., Liang, P., Dean, J., & Fedus, W. (2022, October). Emergent Abilities of Large Language Models [arXiv:2206.07682 [cs]]. Retrieved March 8, 2023, from http://arxiv.org/abs/2206.07682

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D. (2023, January). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models [arXiv:2201.11903 [cs]]. https://doi.org/10.48550/arXiv.2201.11903

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0