Kashihara, A. et al. (Eds.) (2024). Proceedings of the 32nd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Student Perceptions of Using Generative AI Chatbot in Learning Programming

Ean Teng KHOR*, Leta CHAN, Elizabeth KOH & Peter SEOW
National Institute of Education, Nanyang Technological University, Singapore
*eanteng.khor@nie.edu.sg

Abstract: Interest in Generative AI (GenAI) chatbots has exploded across the education industry, subsequently expanding in sophistication and usage. However, students' perceptions on utilizing this technology ultimately determines how well these chatbots promote learning. Hence, this study examines the factors influencing how Singapore secondary school-aged computing students perceive using *MyBotBuddy* (MBB), a GenAI chatbot, to assist them with programming tasks. A thematic analysis determined that students' perceptions were influenced by reliability, utility, cognitive effort needed, satisfaction, and enjoyment. The study contributes to the literature on GenAI chatbots supporting secondary school programming students and may guide the development of such tools in secondary school classrooms more effectively.

Keywords: Generative AI, Chatbot, Computing, Programming, K-12 Students

1. Introduction

Given GenAI's promising potential in supporting learning, global institutions are keen to integrate it into education. Capable of generating customized feedback, providing relevant and concise information, and generating code from natural language, GenAI chatbots promise to improve and personalize programming education (Wu et al., 2023). However, these tools must be smoothly integrated into the classroom for students to reap these promised benefits. To do so, it is critical that students can easily use these chatbots and find them useful, as this determines their will to use GenAI educational tools (Davis, 1989). Despite GenAI's relevance to GenAI chatbots, most educational research on GenAI chatbots focuses on general or medical education and academia (Bahroun et al., 2023) instead. Within GenAI research in this field, higher education is most dominant (Deriba et al., 2024), despite growing interest in K-12 programming education (Szabo et al., 2019). Hence, this study seeks to explore the factors influencing Singapore secondary school computing students' perceptions of using MBB (Khor et al., 2024), a chatbot developed based on the GenAI model, and explore how it can be made easier to use and more useful to them.

2. Research Methods

12 GCE 'O' level computing students, aged 15–16 years old, were recruited from a secondary school in Singapore. Participants were limited to the few consenting students from the small computing class. Consent was also obtained from parents, IRB and the government education ministry. Students' programming knowledge was assessed before and after using MBB with a task to write a Python program calculating the check digit for ISBN-13 (pre-test) and ISBN-10 (post-test) numbers. In between, MBB was introduced as a GenAI chatbot to help improve students' pre-test code or explore better solutions. Computing learning materials trained MBB's knowledge and reasoning abilities. MBB taps on relevant resources, like the GCE 'O' level computing syllabus, to respond appropriately to requests, and can match students' skill based on their grade and provide line by line explanations with worked examples. MBB was

used to explain code, spot errors, de-bug, simplify code, or answer any computing-related questions. Students then completed the post-test without MBB. After the intervention, students answered a survey on their perceptions of and desired functions in educational chatbots, and views on MBB's feedback. As part of a thematic analysis, codes were derived from recurring ideas in key quotes in students' responses. Similar codes were grouped into themes. Using relevant theories, namely Davis' (1989) Technology Acceptance Model (TAM), Menon and Shilpa's (2023) Unified Theory of Acceptance and Use of Technology, Venkatesh and Bala's (2008) TAM3, and Sweller's (1988) cognitive load theory, the themes were sorted as affecting perceived usefulness or perceived ease of use. Sub-themes were identified to reflect how codes related to each theme. A paired *t*-test conducted on the knowledge test scores examined how the themes supported students in practice.

3. Results and Discussion

The survey revealed five themes or factors shaping perceptions of MBB. The themes of reliability and supporting weaker students affected perceived usefulness, while clearly structured information, satisfaction, and enjoyment affected how easy it was to use MBB. Reliability is critical to how useful students found MBB. Students praised the accuracy of MBB's responses (n=2), suggesting that accurate feedback was valued. They also requested to correct logic errors in the generated code (n=2). The conflicting views of MBB's accuracy suggests issues with reliability. Students' requests to correct errors suggest that they may find them detrimental to their learning. While errors were rare, the presence of any may limit MBB's utility, as the generated code would be unusable. How the GenAI model understands and responds to prompts thus needs refining to be reliably accurate. Students also found MBB useful for supporting weaker students. MBB supports the learning of programming concepts (n=8). One of MBB's merits is its step-by-step thorough explanations that even weaker students can understand (S2, S3), and it matches students' pace to enhance their learning (S7). MBB also develops programming skills (n=4) by challenging students' problem-solving skills and improving their coding skills. MBB generates innovative, unique solutions that inspire questions (S7) and challenge how students problem-solve (S4). It also guides students while coding, helping them to identify and learn from their mistakes (S1). However, students noted sparse detail or gaps in MBB's knowledge on some topics (n=2). The chatbot training data thus needs to be equally updated and extensive across topics.

Students' comments on how clearly responses were structured suggest that MBB's design may have influenced its perceived ease of use. One subtheme was the presentation of responses (n=3). The generated responses may be difficult to read, as some suggested highlighting (S1) key words or adding pictures (S7) to improve clarity. Attention needs to be paid to how responses are presented to maximize readability. Another subtheme was the answer length (n=8). Students preferred when answers were concise and clear, with some answers being "very intimidating chunks of text" (S8). Shorter or sectioned responses could make MBB less cognitively taxing to use. Students who found MBB less effortful to understand thus perceived it more positively. Another theme identified was students' satisfaction with MBB (n=4). Students may have found MBB difficult to interact with as they encountered issues with content display capacity (S9) and in understanding prompts (S4). This hindered students from efficiently using MBB, with some expressing displeasure as a result (S7, S9). Attention ought to be paid to the technical features to ensure a satisfying experience to improve its ease of use. Enjoyment, or lack thereof, may also influence students' perceived ease of use of MBB. One student enjoyed exploring the various solutions generated by MBB (S4), and another found MBB "interactive and fun to use" (S10). Learning from a human-like tool may be more engaging and thus enjoyable. Improving these abilities may lead to more enjoyable interactions and reduce the effort needed to use the chatbot.

The study found no significant difference in students' knowledge test scores before (M = 8.2, SD = 2.6) and after (M = 7.8, SD = 3.0) interacting with MBB, $t(11) = .35$, $p = .734$. This is surprising as many found it helpful in supporting their learning (n=7) and beneficial to their programming knowledge (n=9). The knowledge tests and intervention were conducted on the

same day, more interaction time as well as time to consolidate any knowledge learned may have been necessary to see a significant improvement in students' scores. As students noted that MBB is suitable for weaker students, these findings suggest that it may be less useful for skilled students. Hence, it should have equal utility for students of different skill levels when designing a learning tool.

4. Conclusion

Students' perceptions were shaped by reliability, utility in supporting weaker students, satisfaction, enjoyment, and cognitive effort needed to use MBB. Suggestions to enhance perceptions included improving answer accuracy, detail, and scope, sectioning the answers or using visual aids, and improving the model's response speed and comprehension skills. This study also calls for equal support for students regardless of skill level and forms a base for future research on optimizing educational chatbots with students' feedback. Educational stakeholders may use these insights to select well-perceived GenAI tools to promote their acceptance in the classroom. Though the small sample size and focus on a school cohort in Singapore may limit this study's generalizability, it may still help those seeking to improve the general perceptions of AI educational chatbots. Using a survey may also have limited this study's depth, and further exploration to capture students' cognitions and attitudes is needed.

Acknowledgements

This research is supported by the Ministry of Education (MOE), Singapore, under its Academic Research Fund (RG89/22).

References

Bahroun, Z., Anane, C., Ahmed, V., & Zacca, A. (2023). Transforming education: A comprehensive review of generative artificial intelligence in educational settings through bibliometric and content analysis. *Sustainability, 15*(17), 12983. https://doi.org/10.3390/su151712983

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340. https://doi.org/10.2307/249008

Deriba, F., Sanusi, I. T., Campbell, O. O., & Oyelere, S. S. (2024). *Computer programming education in the age of generative AI: Insights from empirical research*. SSRN. https://dx.doi.org/10.2139/ssrn.4891302

Khor, E. T., Koh, E., & Seow, P. (2024, June). Leveraging generative artificial Intelligence to support students in learning programming: A chatbot pilot study. In *Proceedings of the 18th International Conference on Learning Sciences*. International Society of the Learning Sciences. United States: Buffalo.

Menon, D., & Shilpa, K. (2023). "Chatting with ChatGPT": Analyzing the factors influencing users' intention to use the OpenAI's ChatGPT using the UTAUT model. *Heliyon, 9*, Article e20962. https://doi.org/10.1016/j.heliyon.2023.e20962

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257-285. https://doi.org/10.1207/s15516709cog1202_4

Szabo, C., Sheard, J., Luxton-Reilly, A., Simon, Becker, B. A., & Ott, L. (2019, November). Fifteen years of introductory programming in schools: a global overview of K-12 initiatives. In *Proceedings of the 19th Koli Calling International Conference on Computing Education Research* (pp. 1-9).

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences, 39*(2), 273-315. https://doi.org/10.1111/j.1540-5915.2008.00192.x

Wu, T. T., Lee, H. Y., Li, P. H., Huang, C. N., & Huang, Y. M. (2023). Promoting self-regulation progress and knowledge construction in blended learning via ChatGPT-based learning aid. *Journal of Educational Computing Research, 61*(8), 3 - 31. https://doi.org/10.1177/07356331231191125