Assessing the Influence of ChatGPT on Student Outcomes in a Models of Computing Course

Iñaki Blasco
University of British Columbia
Vancouver, Canada
ikblasco@student.ubc.ca

Karina Mochetti
University of British Columbia
Vancouver, Canada
mochetti@cs.ubc.ca

Abstract

This study investigates the impact of ChatGPT on student performance in a Models of Computing course, foundational for the computer science major. Analysing data from 11 pre-lecture quizzes across four terms, we found a decline in average quiz scores, particularly in the latest term. The results suggest a correlation between increased reliance on ChatGPT and decreased student performance, especially on challenging questions where the AI frequently struggled. These findings highlight both the benefits and challenges of integrating AI in education. Our ongoing research aims to explore this further across multiple courses, ultimately promoting responsible AI use to enhance learning outcomes.

CCS Concepts

• **Social and professional topics** → *Computing education*; • **Theory of computation** → *Models of computation*; • **Computing methodologies** → *Artificial intelligence*.

Keywords

Computing education, LLM, Student Performance

ACM Reference Format:
Iñaki Blasco and Karina Mochetti. 2025. Assessing the Influence of ChatGPT on Student Outcomes in a Models of Computing Course. In *Proceedings of the 56th ACM Technical Symposium on Computer Science Education V. 2 (SIGCSE TS 2025), February 26-March 1, 2025, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3641555.3705132

1 Introduction and Motivation

In recent years, artificial intelligence (AI) has transformed many aspects of human life, influencing industries such as healthcare, entertainment, and education. Among the most significant advancements in this field is ChatGPT, developed by OpenAI [5]. This powerful language model utilizes deep learning to generate human-like text based on user prompts, drawing considerable attention for its ability to understand and produce coherent, contextually relevant language.

While ChatGPT offers numerous benefits in education [2, 3], it also presents potential challenges, particularly regarding academic integrity. The accessibility of sophisticated AI-driven assistance can tempt some students to misuse the technology for dishonest purposes. Such misuse undermines the educational process, as it enables students to bypass critical learning experiences and fail to develop essential problem-solving skills [1, 4]. Furthermore, it poses a significant challenge for educators, who must maintain academic standards and ensure that assessments accurately reflect students' understanding and abilities. As AI tools like ChatGPT become increasingly integrated into educational settings, it is crucial to establish clear guidelines and implement robust detection mechanisms to mitigate the risks of academic dishonesty.

This research project aims to examine the impact of ChatGPT on student performance, focusing on its role in pre-lecture quizzes within the first-year Computer Science course, Models of Computation (CPSC 121) from UBC. Through a detailed case study, we will analyse ChatGPT's responses and compare them with actual student outcomes to assess its accuracy. Our goal is to provide clear data on how ChatGPT performs in this course. We plan to use the results to foster discussions with students on the importance of critical thinking and the responsible use of available tools to support their learning.

2 Methodology

CPSC 121 is a foundational Computer Science course that covers essential math and logic concepts for a career in computer science. Key topics include proofs, formal logic, circuits, deterministic finite automata (DFAs), and discrete mathematics.

The primary data source for our analysis will be the averages of pre-lecture quizzes. These quizzes are assigned to students after they complete the required reading and before attending class at the start of each new module. Each quiz consists of nine multiple-choice questions and one open-ended question, designed to assess students' foundational knowledge prior to lectures. The purpose of these quizzes is to prepare students for fast-paced lectures by ensuring they arrive with a solid understanding of the material and relevant questions. These quizzes are consistent each term, allowing students two attempts at each to encourage mastery of the content.

The data timeline spans from term 2022W1 to 2023W2, covering two academic years, each consisting of two semesters. The first semester runs from September to December, while the second semester spans January to May. This timeline enables us to account for the evolution of ChatGPT models and how students may have used different versions over time.

We compare the average of each questions to how well ChatGPT performs in answering it in different versions and formats. Due to updates and subscription limitations, we primarily utilized versions GPT-3.5, GPT-4, and GPT-4o. We tested the model using various

SIGCSE TS 2025, February 26-March 1, 2025, Pittsburgh, PA, USA

ACM ISBN 979-8-4007-0532-8/25/02
https://doi.org/10.1145/3641555.3705132

formats multiple times to simulate how students might interact with ChatGPT for quiz assistance. The three formats tested were:

- **Format 1:** Submitting all quiz PDFs to ChatGPT simultaneously and requesting answers (not applicable to GPT-3.5).
- **Format 2:** Submitting each quiz PDF independently to ChatGPT and requesting answers (not applicable to GPT-3.5).
- **Format 3:** Manually typing each question into ChatGPT and adding images when possible.

These formats reflect potential variations in how students might pose questions at different times. Our goal was to simulate a realistic student experience when seeking AI assistance.

To analyse the impact of ChatGPT on student performance, we compared quiz averages across different terms with Chat GPT average performance. By analysing the consistency and accuracy of answers provided by each model, we aimed to identify trends and patterns that suggest how advancements in AI might influence student learning outcomes. Additionally, through patterns in the data, we seek to highlight the usage of ChatGPT among students.

3 Results and Impact

We analysed all 11 quizzes taken in CPSC 121 over the course of four terms. Figures 1 and 2 illustrate the average scores for Quiz 02 and Quiz 08 during this period, alongside the performance of ChatGPT on each question. The background color indicates how well ChatGPT performed on each question, based on the various versions and formats we tested.

Notably, the average quiz scores have decreased over time, with Term 4 (2023W2) showing the lowest performance. Questions that experienced the most significant declines typically corresponded to those where ChatGPT often provided incorrect answers.

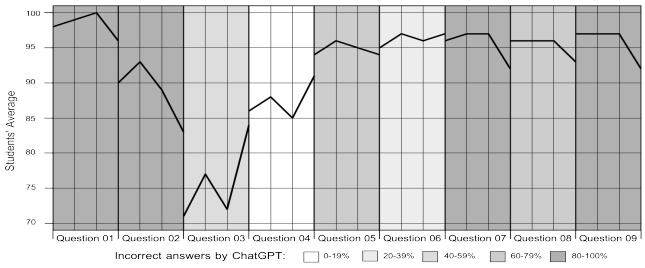

Figure 1: Average of each question for Quiz 02 during 4 different terms compared with ChatGPT performance.

For Quiz 02, questions 3 and 4, which ChatGPT answered correctly most of the time, showed an increase in average scores. Conversely, questions that ChatGPT frequently answered incorrectly experienced a significant decline, especially in the last term analysed, which occurred in early 2024 when ChatGPT was commonly used. This trend is particularly evident in more challenging questions, like question 3, where students might have been more inclined to seek help from ChatGPT.

A similar pattern emerged in Quiz 08. The last question, the most difficult in that quiz, demonstrated a noticeable decrease in average scores, with ChatGPT failing to answer it correctly most of the time. Question 5, which also posed challenges for ChatGPT, saw a decline

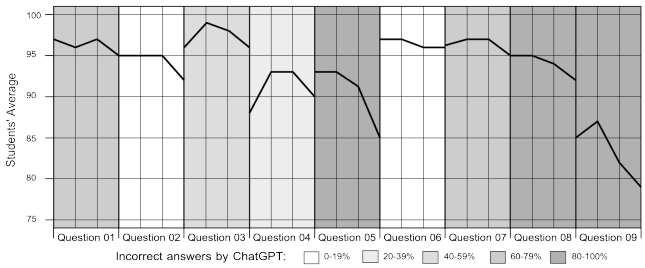

Figure 2: Average of each question for Quiz 08 during 4 different terms compared with ChatGPT performance.

in scores as well. In contrast, easier questions, such as questions 1 and 3, where ChatGPT performed well and most students also succeeded, showed little variation between terms. This suggests that students likely sought ChatGPT's assistance primarily for the more difficult questions.

This pattern has been observed across all quizzes in CPSC 121, with these two serving as clear examples. By analysing data from pre-lecture quizzes in CPSC 121, we identified significant trends that highlight both the potential benefits and challenges of integrating AI into education.

Preliminary results indicate that average quiz scores declined over the analysed terms, with the most substantial drops occurring in the later terms. This trend suggests a possible correlation between increased ChatGPT usage and decreased student performance.

Our next steps involve conducting long-term studies to track the evolution of student performance over multiple academic years, which will provide a more comprehensive understanding of how AI tools influence learning outcomes over time, especially considering the improvements in AI technology. We also aim to analyse which topics and types of questions are most affected by this trend.

Recognizing that banning ChatGPT is not a viable solution, our primary objective is to provide a framework for both students and educators on the appropriate use of AI tools. This will help maximize their benefits while minimizing potential drawbacks. A key goal is to present clear data to students and foster discussions on the careful use of ChatGPT, emphasizing the need for critical thinking. Students should understand that, while ChatGPT can be a valuable resource, it is not a substitute for effective tutoring if used carelessly.

References

[1] Bruno Pereira Cipriano and Pedro Alves. 2024. "ChatGPT Is Here to Help, Not to Replace Anybody"–An Evaluation of Students' Opinions On Integrating ChatGPT In CS Courses. *arXiv preprint arXiv:2404.17443* (2024).

[2] F Abu Deeb and Timothy Hickey. 2015. Spinoza: the code tutor. In *International Conference on Computer and Information Science and Technology*.

[3] Barbara J Ericson, Paul Denny, James Prather, Rodrigo Duran, Arto Hellas, Juho Leinonen, Craig S Miller, Briana B Morrison, Janice L Pearce, and Susan H Rodger. 2022. Parsons problems and beyond: Systematic literature review and empirical study designs. *Proceedings of the 2022 Working Group Reports on Innovation and Technology in Computer Science Education* (2022), 191–234.

[4] Md Mostafizer Rahman and Yutaka Watanobe. 2023. ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences* 13, 9 (2023), 5783.

[5] Tianyu Wu, Shizhu He, Jingping Liu, Siqi Sun, Kang Liu, Qing-Long Han, and Yang Tang. 2023. A brief overview of ChatGPT: The history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica* 10, 5 (2023), 1122–1136.

Temporary page!

LaTeX was unable to guess the total number of pages correctly. As there was some unprocessed data that should have been added to the final page this extra page has been added to receive it.

If you rerun the document (without altering it) this surplus page will go away, because LaTeX now knows how many pages to expect for this document.