www.ijemst.net

ChatGPT as a Learning Assistant: Its Impact on Students Learning and Experiences

Joel B. Jalon, Jr.
De La Salle University - Integrated School, Philippines

Goodwin A Chua
De La Salle University - Integrated School, Philippines

Myrla de Luna Torres
De La Salle University - Integrated School, Philippines

To cite this article:

Jalon, J.B., Chua, G.A., & Torres, M.L. (2024). ChatGPT as a learning assistant: Its impact on students learning and experiences. *International Journal of Education in Mathematics, Science, and Technology (IJEMST), 12*(6), 1603-1619. https://doi.org/10.46328/ijemst.4471

International Journal of Education in Mathematics, Science and Technology

2024, Vol. 12, No. 6, 1603-1619 https://doi.org/10.46328/ijemst.4471

ChatGPT as a Learning Assistant: Its Impact on Students Learning and Experiences

Joel B. Jalon, Goodwin A Chua, Myrla de Luna Torres

Article Info

Article History

Received:

03 May 2024

Accepted:

07 September 2024

Keywords

ChatGPT

Learning assistant

Python

Programming tutor

Abstract

ChatGPT is largely acknowledged for its substantial capacity to enhance the teaching and learning process despite some concerns. Based on the available literature, no study compares groups of students using ChatGPT and those who did not, more so in programming. Therefore, the main goal of this study was to examine how ChatGPT affects SHS students' learning of Python programming with students who did not use the AI tool. An independent t-test revealed that there is a significant difference between students who utilized ChatGPT and high-performing students who did not use it, but the observed difference is not likely to have practical significance, indicating the potential of ChatGPT to enhance students' programming proficiency. The effectiveness of ChatGPT in providing solutions and improving understanding was also supported through students' experiences. Four themes were derived from the thematic analysis of students' shared experiences: students were appreciative and satisfied, students were amazed and surprised by ChatGPT's ability, students acknowledged ChatGPT's limitations, and students treated ChatGPT like a human being. The study concludes that ChatGPT can complement conventional teaching approaches but underscores the importance of independently cultivating students' problem-solving abilities, and cooperation between educators and developers can augment the role of AI-driven tools in education.

Introduction

Artificial intelligence, or AI, a product of technological advancement, was created to aid creatively challenged people and produce textual materials and programs, leading to AI-powered writing assistants' emergence. AI significantly affects people's lives. Hence, education stakeholders should be prepared to adapt to this revolution (Pokkakillath & Suleri, 2023). Chat Generative Pretrained Transformer 3, or ChatGPT-3, was released in November 2022 by OpenAI, which attracted the attention of millions and made ChatGPT a headline (Ding et al., 2023; Kiryakova & Angelova, 2023; Pokkakillath & Suleri, 2023). A substantial amount of data is used to train ChatGPT, which OpenAI has yet to publicize this information (Stojanov, 2023). Since ChatGPT can produce written text that closely resembles human language, one of the most significant discussions is around its use in

education and academia, and its potential has spurred debates about its consequences (Tlili et al., 2023; Wardat et al., 2023). It attracted much student interest and showed potential as a valuable tool to assist students in their educational journeys (Ngo, 2023). Rejecting or forbidding it will not lessen its impact (Wardat et al., 2023). Hence, banning ChatGPT because of ignorance or legal restrictions is impractical and needs more innovation, considering its significant utility as a tool (Popovici, 2023). A similar viewpoint was suggested by Tlili et al. (2023), that instead of banning ChatGPT, embrace it, adopt a new view of technology, be critical of ChatGPT's responses, and identify and improve skills that can affect ChatGPT's responses. On the one hand, the developers of large language models should consider humanized chatbots and establish guidelines consistent with the law and human values, in addition to privacy and security.

Pokkakillath and Suleri (2023) recommended investing in research to explore the effect of ChatGPT on learners and the education system more. Furthermore, more studies are needed on integrating ChatGPT in teaching programming, resulting in more data to deduce conclusions (Husain, 2024). Existing studies are more about education stakeholders' feedback (Husain, 2024; Kiryakova & Angelova, 2023; Ngo, 2023; Pokkakillath & Suleri, 2023; Teubner et al., 2023; Tlili et al., 2023; Zhu et al., 2023), and no study compares groups of students using ChatGPT with those who used it (Ding et al., 2023; Jauhiainen & Guerra, 2023; Liang et al., 2023; Stojanov, 2023; Wardat et al., 2023), more so in programming (Coello et al., 2024; Hartley et al., 2024; Jalil et al., 2023; Jin & Kim, 2023; Jing et al., 2024; Popovici, 2023). Therefore, the main goal of this study was to examine how ChatGPT affects SHS students' learning of Python programming with students who did not use the AI tool. The study looked at students' programming proficiency and experiences using ChatGPT as a teaching tool to shed light on the function of AI-driven tools in programming instruction. Specifically, this study aimed to answer the following research questions.

1. How does using ChatGPT affect students' proficiency in Python programming?
2. What are the students' experiences using ChatGPT to learn to program?

Based on the available literature, no study compares groups of students using ChatGPT and those who did not, more so in programming. Through analyzing exam scores and student experiences, this study offers empirical evidence to guide decision-making processes regarding incorporating ChatGPT or comparable technologies into instruction programming. The result of this study can provide valuable insights into how ChatGPT might be effectively utilized in programming courses while shedding light on its possible advantages and drawbacks. Teachers can utilize this information to create more efficient educational experiences that suitably incorporate AI tools such as ChatGPT. This study provides valuable insights for advancing and enhancing AI-powered educational systems such as ChatGPT. This research enhances the continuing development of technologies that promote student learning by identifying areas of proficiency and areas for improvement.

Literature Review

Large Language Models and ChatGPT

Large Laguange Models (LLMs) mix massive volumes of textual training data with large-scale architectures. Due to this scaling up, LLMs can now produce and comprehend text at a level like that of humans (Teubner et al.,

2023). Some known LLMs are Bard, Bing, Claude, and ChatGPT. OpenAI (2022) defines ChatGPT as an online AI-powered platform. It performs several functions: composition, summarization, and text arrangement. ChatGPT is designed to interpret and respond to user queries and commands. As an LMM, it uses a large amount of textual data to provide contextual comprehension. It uses predictive modeling, which anticipates and generates the following words based on learning patterns, like auto-complete functions in email systems, smartphones, and search engines. According to OpenAI (2022), ChatGPT was trained using feedback and human conversations. In the beginning, they had people role-play dialogues between themselves and an AI assistant while having access to the assistant's recommendations. Subsequently, they collected data in which persons evaluated various AI answers according to their suitability. They improved ChatGPT's performance using this data and a technique known as Proximal Policy Optimization. To make the model better, they went through this process multiple times.

ChatGPT in Education

The adoption of Language Model Models (LLMs) like ChatGPT in the field of education is inevitable. Pokkakillath & Suleri (2023) listed how educational institutions use AI, such as generating learning materials according to the learners' learning styles and interests, checking and giving feedback on student's essays, creating assessments, virtual personal tutoring, and designing teaching materials, composing scholarly articles. According to Wardat et al. (2023), chatbots can be used in the classroom to evolve an educational philosophy. It also emphasizes how important it is to update competencies to help teachers and students become more adept at using chatbot technology.

ChatGPT has the power to revolutionize education by providing many advantages that could help teachers and students in different ways, such as designing instructional materials and creating tests (Tlili et al., 2023). Jalil et al. (2023) also proposed that ChatGPT can be a guiding tool, rather than a conventional teacher or teaching assistant, to enhance students' comprehension during in-class activities or laboratory sessions. Based on the research conducted by Tlili et al. (2023), Ngo (2023), and Husain (2024) on the educational stakeholder's views, there is a widespread agreement that ChatGPT is generally regarded positively. Users value its user-friendly interface, convenience, and ability to transform education by providing tailored learning opportunities. Moreover, in the study of Ngo (2023), users recognize the effectiveness of ChatGPT in enhancing knowledge acquisition, maximizing time utilization, and providing access to an immense amount of information. ChatGPT is widely recognized as a good tool for improving memory and academic performance, particularly among college students.

Several researchers personally evaluated ChatGPT across several subjects. These studies have noted that ChatGPT's responses demonstrated varying degrees of precision and dependability across different subjects. While ChatGPT demonstrates proficiency in certain areas, such as solving physics problems in motion (Liang et al., 2023), its ability to handle software testing questions could be more consistent (Jalil et al., 2023). Although it can offer precise solutions to many queries, there are ongoing concerns about its tendency to generate incorrect answers, especially when faced with complex questions that involve both coding and conceptual elements. In addition, ChatGPT's comprehension of geometry is limited, and it cannot rectify misconceptions proficiently (Wardat et al., 2023).

Popovici (2023) and Coello et al. (2024) also had similar results: ChatGPT generated incorrect codes. Nonetheless, Coello et al. (2024) demonstrated the exceptional proficiency displayed by ChatGPT in generating Python programs and its capacity to generate more efficient code than other Language Model-based Models (LMMs). Despite some limitations, such as the inconsistency in its generated responses and potential readability concerns, ChatGPT significantly assists students in effectively acquiring fundamental programming skills by generating functional code.

Jauhiainen and Guerra (2023) and Jin and Kim (2023) included ChatGPT in developing educational materials. Jauhiainen and Guerra (2023) claimed that ChatGPT can tailor educational materials to match learners' cognitive abilities. ChatGPT can instantly generate and edit information, text, figures, and tasks. This allows the system to customize instructional materials according to each learner's needs and preferences. These findings are corroborated by Jin and Kim (2023), who reported that educators and developers expressed favorable opinions regarding using ChatGPT to customize educational resources in the instruction of programming languages. Moreover, the incorporation of ChatGPT has resulted in substantial enhancements in learners' comprehension of the programming language. The research indicates that ChatGPT can enhance learning outcomes in programming education.

ChatGPT does not only affect students' cognitive development but also the affective domain. Liang et al. (2023) and Jin and Kim (2023) discovered that ChatGPT's human-like interaction favors learners' motivation. The tool's interactive discussion could enhance students' motivation and interest in the subject matter. Even with its benefits, users have voiced concerns regarding the constraints of ChatGPT. According to the research conducted by Kiryakova and Angelova (2023), professors are hesitant to incorporate ChatGPT into their classes, unlike college students, who generally embrace it. The professors' unwillingness to accept integrated ChatGPT in their instruction is driven by apprehensions about its lack of precision and the potential adverse effects on analytical thinking and innovation. Similar issues were also raised by Ngo (2023) and Husain (2024). More researchers (Pokkakillath & Suleri, 2023; Tlili et al., 2023; Zhu et al., 2023) concurred that an overreliance on ChatGPT can harm one's critical thinking skills.

Excessive dependence on ChatGPT for instant answers might gradually erode students' ability to think critically and make decisions, leading to a decline in these skills. Concerns regarding the precision and dependability of information generated by ChatGPT can offer challenges for students who rely on it for educational and research endeavors. There are also concerns about the potential for plagiarism, particularly if students use ChatGPT to generate content without giving proper credit (Pokkakillath & Suleri, 2023; Tlili et al., 2023; Popovici, 2023).

Researchers emphasized the importance of verifying ChatGPT's responses rather than unquestioningly accepting them as absolute answers, especially as the complexity of educational tasks increases (Popovici, 2023; Wardat et al., 2023). This recommendation was also supported by the findings of Ding et al. (2023). While ChatGPT can correct its answers based on user prompts, there remains a significant concern over the potential for generating erroneous responses (Zhu et al., 2023). Rather than banning ChatGPT, teachers should teach students how to use it (Popovici, 2023; Tlili et al., 2023; Wardat et al., 2023). The effectiveness of incorporating ChatGPT in various

educational fields is influenced by different factors and approaches, as evidenced by the studies conducted by Jing et al. (2024) and Zhu et al. (2023).

Jing et al. (2024) highlighted the significance of AI literacy in evaluating the effectiveness of using ChatGPT for solving programming problems. The level of learners' knowledge and application of AI, the aptitude of learners in Python, and their understanding of ChatGPT significantly impact their ability to utilize ChatGPT effectively, but not the learners' willingness to use ChatGPT, the use of AI assessment, and adherence to AI ethical principles. Furthermore, the study emphasizes the importance of adopting a fair and balanced attitude while dealing with new technology, avoiding overly cautious and overly optimistic views to maximize educational benefits. On the other hand, Zhu et al. (2023) provided methods to improve the efficiency of ChatGPT in educational settings, employing the Content, Others, Self, and Tasks (COST) model framework. This approach proposes employing ChatGPT as a tool that helps create content materials, a virtual instructor and study companion, a tool that promotes reflective learning, assistance in completing tasks, and a catalyst for fostering critical thinking.

Theoretical Framework

This study is guided by social constructivism by Vygotsky (1979). The central idea of social constructivism is that there are two stages in a learner's construction of knowledge: the learner's interaction with others and the learner's self-realization from his interaction. Furthermore, higher cognitive processes such as analyzing and evaluating are learned through the learner's interaction with people. Social constructivism stresses the collaborative nature of learning. This idea differs from Piaget's Constructivism, which says that a learner constructs new knowledge and gives meaning based on the learner's experience.

According to Vygotsky (1979), language affects how learners think, and as language develops, so does one's thinking. For children, signs and words are ways to interact with other people, which becomes the primary method in their interaction, differentiating them from animals. Thus, language is essential. Furthermore, Vygotsky (1979) introduced the zone of proximal development, *the distance between the actual development level determined by independent problem-solving and the level of potential development as determined through problem-solving under adult guidance or in collaboration with more capable peers*. The developmental process is slower than the learning process, creating the proximal development zone. However, a learner has the potential to pass through and cross the zone of proximal development and acquire knowledge and skills that cannot be acquired by himself with the aid of a person who is more knowledgeable and more skillful, whom the learner can observe and practice his skills and by providing instructional scaffolding which guides the learner through the zone of proximal development. According to Stojanov (2023), this more knowledgeable other could be written text, videos, computers, the internet, and even AI such as ChatGPT, which are not humans.

ChatGPT's versatility, interaction, and proficiency in offering tailored assistance and feedback make it a potential tool for tutoring and aiding student learning in diverse disciplines and situations. It provides customized learning experiences through personalized programming activities, coding assistance, and practice opportunities to meet each student's unique needs and learning preferences (Jauhiainen & Guerra, 2023; Jin & Kim, 2023; Kiryakova

& Angelova, 2023). It can provide fast feedback and aid students in programming (Coello et al., 2024; Hartley et al., 2024; Jalil et al., 2023; Jin & Kim, 2023; Jing et al., 2024; Popovici, 2023).

Method

Study Design

This study utilized a quasi-experimental research approach. In a quasi-experimental approach, the researcher needs more control over the distribution of participants to various groups. Alternatively, the researcher relies on naturally existing groups or circumstances, such as students who voluntarily opt to utilize ChatGPT compared to those who abstain. Although the researcher is unable to employ random assignment of participants to groups, they can still conduct a comparative analysis of outcomes across the groups and make inferences on the likely impact of the independent variable (in this instance, ChatGPT usage) on the dependent variables (exam performance and user's experiences) (Creswell & Creswell, 2018).

This study also utilized a mixed-method approach, incorporating quantitative and qualitative data (Creswell & Creswell, 2018). The students' final exam scores were examined using an independent t-test to identify if there was a significant difference between the test scores of the students who used ChatGPT and those who did not to gain an understanding of the influence of ChatGPT on their academic achievement. Furthermore, the participants provided feedback on their interactions with ChatGPT by answering an online survey form, which was then subjected to thematic analysis.

ChatGPT was embedded in the teaching strategy for all the classes taking Empowerment Technologies 1 (EMTECH1). Each class was divided into groups of two to three members to work on two programming problem sets and the course's term project. This was to lessen their workload on the subject's requirements. The teacher announced to the students that they could use ChatGPT to answer the two problem sets on conditional and iterative statements and their term project, which required them to create Python programs, provided that they had tried to answer the tasks first. Hence, the students were not required to use ChatGPT if they did not need it. This is to ensure that students use their critical thinking and problem-solving skills, as AI tools may negatively affect learners' critical-thinking skills (Tlili et al., 2023; Pokkakillath & Suleri, 2023; Zhu et al., 2023), but students can develop critical thinking skills through discussion with ChatGPT (Zhu et al., 2023). The teacher set one session, 1.5 hours long, per problem set, while the term project was done during the class for five sessions and outside class hours.

Participants

The participants of this study were 82 grade 11 senior high school students of De La Salle University - Integrated School taking EMTECH1 from three intact classes; 32 belong to one Science, Technology, Engineering, and Mathematics or STEM class, and 50 belong to two Art and Design Track or (ADT) classes. EMTECH1 tackles computational thinking concepts and basic Python programming. Python lessons include variables, input, output, conditional, and iterative statements (While loop only).

Instruments

Programming Tasks

There were two problem sets to test students' conceptual understanding and programming skills: one for Conditional Statements and another one for Iterative Statements. These problem sets were administered via HackerRank, an online programming platform with an auto-check feature. Hence, the students knew if their answers were correct or if they needed to change their answers. The Conditional Statements Problem Set has seven easy items, while the Iterative Statements Problem Set contains 14 items ranging from easy to difficult. Students were asked to answer a pre-identified item for each problem set, in which they created a tree diagram, a flowchart, and a Python program. The students were also asked to choose and answer two more items for each problem set, but only Python programs were required. Figure 1 shows examples of items for the problem sets.

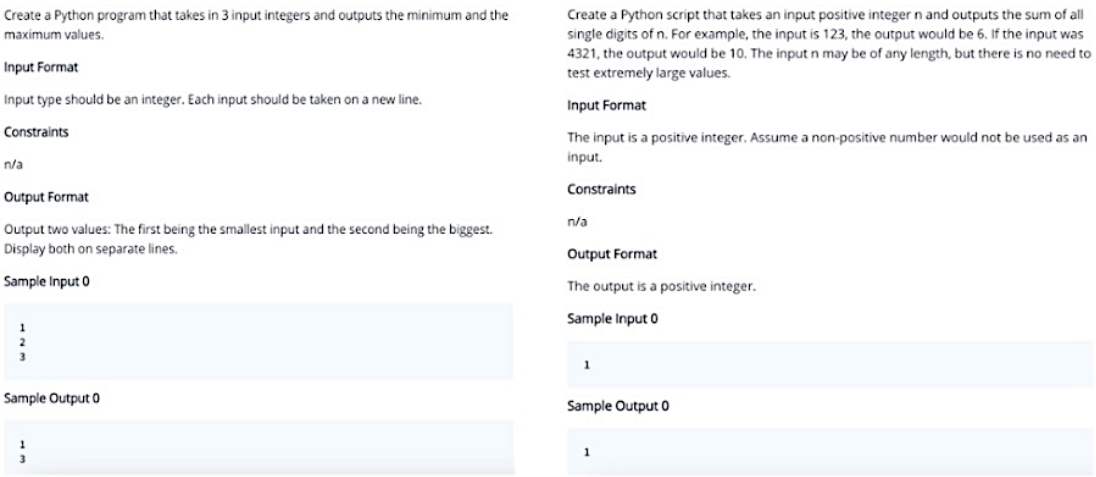


Figure 1. Sample Items for the Problem Sets

Term Project

The term project was announced at least a month before its deadline. The students were asked to create a simple cash management system. The system should allow users to perform various financial transactions, such as cash in, cash out, transfers, and bill payments. The system should also keep track of the user's transaction history. There were restrictions to prevent the students from using AI to generate the entire program. The students were not allowed to use commands or codes that were not part of the discussion, such as using for loops, creating functions, using "True," "False," "None," "continue," "break," arrays, lists, tuples, and importing a library. Students can work alone or with a maximum of two classmates to avoid social loafing. They worked outside the class hours independently and were given at least five class sessions to do the term project.

ChatGPT 3.5

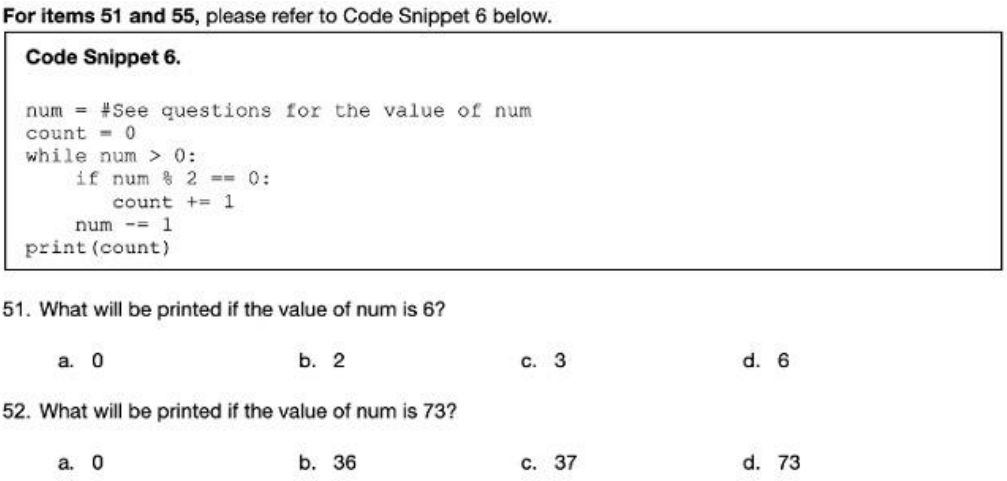
The students were given the option to use ChatGPT. ChatGPT was chosen because it is the most familiar LMM to students. It comes second only behind ChatGPT 4 in producing Python programs and passing assertion tests (Coello et al., 2024). Moreover, it provides accessibility without the need for a subscription.

Student's Feedback Questionnaire

The Student's Feedback Questionnaire (SFQ) had two parts. The first part was to collect the participants' consent and demographic profiles such as sex, age, and grade level. The second part contains questions about whether they already used ChatGPT before and which programming tasks they used ChatGPT, an open-ended question asking the students to share their experience, and an optional field for the participants to upload photos of their chat with chatGPT. The open-ended question was, "Please share your experience using ChatGPT to answer any of the programming tasks." The SFQ was presented to the other five teachers teaching EMTECH1 for feedback and possible improvements to establish validity. The SFQ was administered to the students through a GForm at the end of the course before the final exam.

Final Exam

There were two sets for the final exam containing 58 multiple-choice items. The items in Set A corresponded to those in Set B, meaning they were parallel. The final exam items were reviewed by eight teachers teaching EMTECH1 and were used for two consecutive terms. Figure 2 shows example items for the final exam. Only the last 28 items, which were about programming, were included in the analysis. The internal reliability indices of the two sets with the 28 items were computed using Cronbach's α. Set A has a reliability index of 0.877, and Set B has a reliability index of 0.866. These results suggested that the two sets have good reliability.



For items 51 and 55, please refer to Code Snippet 6 below.

Code Snippet 6.

```
num = #See questions for the value of num
count = 0
while num > 0:
    if num % 2 == 0:
       count += 1
    num -= 1
print(count)
```

51. What will be printed if the value of num is 6?

a. 0 b. 2 c. 3 d. 6

52. What will be printed if the value of num is 73?

a. 0 b. 36 c. 37 d. 73

Figure 2. Sample Final Exam Items

Data Collection

At the end of the term, consent was obtained from the students to gather and use their data for this study. Hence, only those who agreed to participate answered the SFQ and included their final exam scores for this study. To answer the first research question, "How does using ChatGPT affect students' proficiency in Python programming?" the final exam scores of the students who used ChatGPT and those who did not in the 28 items were subjected to normality test via the Shapiro-Wilk Test and homogeneity of variances test via the Levene's test. It was found that an independent t-test was applicable to determine if there was a significant difference

between the final exam mean scores.

To answer the second research question, “What are the students' experiences using ChatGPT to learn to program?” students' answers to the open-ended question in the SFQ were subjected to thematic analysis. As Braun and Clarke (2006) described, thematic analysis is used to identify, analyze, and interpret essential patterns, referred to as "themes," in qualitative data. It systematically generates codes, using the smallest units that contain pertinent data pieces that combine to form bigger patterns known as themes. These topics serve as a structure for organizing and showcasing analytical observations, all guided by a central concept. The thematic analysis aims to identify and analyze the key elements of the data while adapting to the evolving research question. Thematic analysis can identify patterns in the data and establish links to the participants' experiences, attitudes, and behaviors. This study adopted the phases suggested by Braun and Clarke (2006): familiarizing with the data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the report.

Moreover, to establish reliability for each phase, the researchers adopted recommendations by Nowel et al. (2017), as described below. Two researchers headed the thematic analysis. In Phase 1, the researchers reviewed the data at least two times to get familiar with all aspects of the data. In Phase 2, the researchers compared their notes and the codes they generated for triangulation, looking for similarities and further discussing the differences. In Phase 3, the researchers independently employed an inductive methodology to identify the themes. In an inductive approach, the identified themes are directly linked to the data and may not necessarily correspond precisely to the precise questions asked of the participants. Inductive analysis entails coding the data without fitting it into a prior coding framework or the researcher's predetermined analytical ideas (Creswell & Creswell, 2018). In Phases 4 and 5, the researchers presented the themes to each other. They deliberated on whether the identified themes accurately represented the meanings discovered in the data set and were sufficiently distinct and comprehensive to justify discontinuing further revisions. In Phase 6, the researchers integrated the participants' quotes into the report and connected the findings to available literature to strengthen the contextual description. In addition, the researchers progressed from providing a simple description to offering an interpretation during the analytical process. The researchers carefully evaluated the importance of the found patterns, reflecting on their wider significance and consequences concerning the research questions.

Results

Implementing ChatGPT as an educational tool for Python programming among students produced fascinating outcomes, revealing insights into its benefits and constraints. This discussion part explores the results of the research questions, providing insights into the influence of ChatGPT on students' programming proficiency and overall experiences with the tool.

Respondents

Twenty-nine (94%) of the groups responded to the survey. Moreover, 21 (72%) groups used ChatGPT to answer any programming tasks. Most groups (19 or 44%) used ChatGPT to answer the problem set in iterative statements.

This observation was because students found it challenging to comprehend iterative statements; the teacher extended the discussion to another meeting for this lesson before the students answered the problem set.

Twelve (28%) groups needed ChatGPT's help to answer the conditional statements. Only four (9%) groups needed ChatGPT's assistance in creating the program for the term project. This result can be ascribed to the limits set by the teachers to discourage excessive dependence on AI-generated solutions. Efforts to produce the entire program using ChatGPT were impeded by its failure to precisely adhere to the given prompts, thereby revealing ChatGPT's constraint. The researchers tried to generate the entire program using ChatGPT 3.5, and even if the prompt was specific not to use lists, for loops, and create functions, ChatGPT 3.5 still includes these in the program, as shown in Figure 3. Follow-up prompts were ineffective. Hence, it would be impossible for the students to use ChatGPT to generate the entire program. This is consistent with the study of Popovici (2023), which states that as the complexity of the learning tasks increases, ChatGPT cannot handle the demands efficiently and depends solely on its answers, as final instructions may often be counterproductive. This ChatGPT's constrain can be used as an advantage that can address the overreliance of students on ChatGPT (Pokkakillath & Suleri, 2023; Tlili et al., 2023; Zhu et al., 2023) in programming courses.

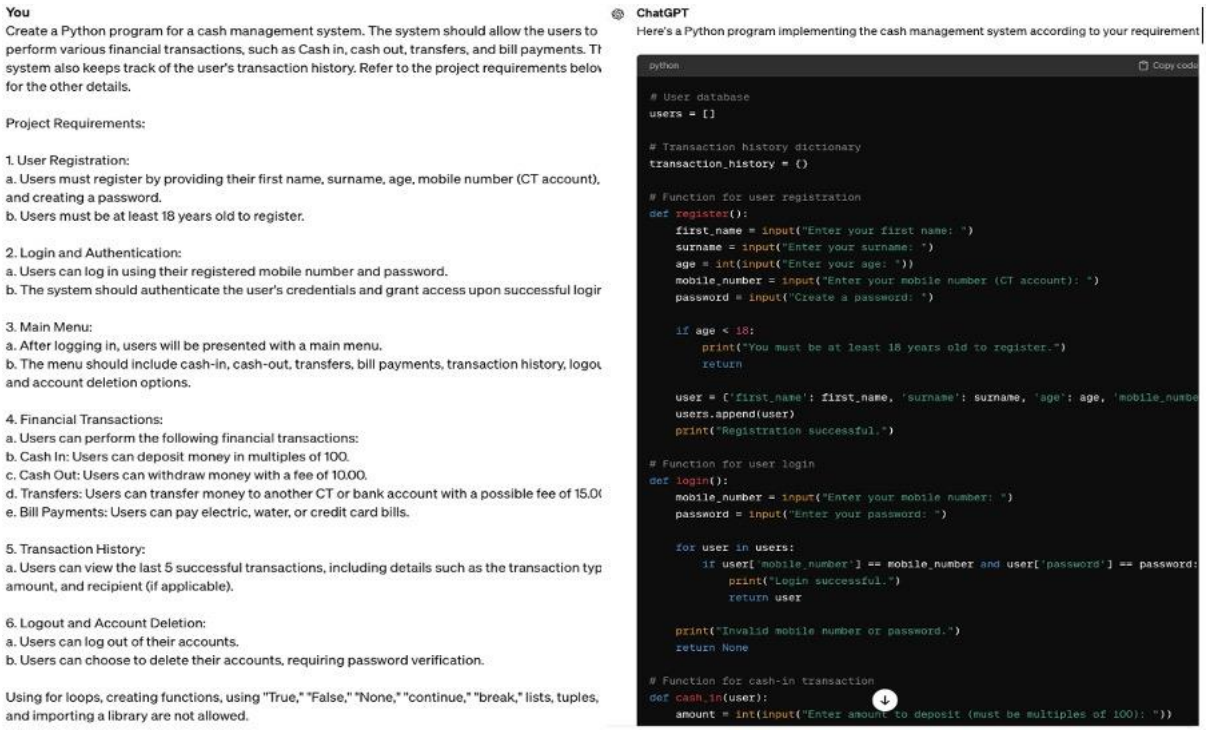


Figure 3. ChatGPT 3.5's Answer to the Term Project Based on a Prompt

RQ1: How does using ChatGPT affect students' proficiency in Python programming?

To answer the first research question, the scores of the students in the final exam were analyzed. Only the last 28 items that were about programming were included in the analysis. Of 82 participants, 64 (79%) agreed to collect and use their data for this survey. The participants were 16 to 18 years old (x=16.64). In addition, most respondents (57 or 89%) had used ChatGPT way back.

Table 1. Descriptive Statistics of the Scores Split by Group

Group	N	Mean	SD	Min	Max	Shapiro-Wilk W	Shapiro-Wilk p
Didn't use	15	20.2	4.62	12	27	0.937	0.346
Used ChatGPT	49	18.5	5.28	8	28	0.975	0.389

Forty-nine (77%) students from the 21 groups used ChatGPT to answer any programming tasks, and 15 (23%) students from the eight groups did not need ChatGPT's assistance. Students' exam scores who used ChatGPT ranged from 8 to 28 (M=18.5, SD=5.28), while those who did not scored 12 to 27 (M=20.2, SD=4.62). Furthermore, the Shapiro-Wilk test showed that both scores were normally distributed: Did not use (*W=27.0, p=0.937*) and Used ChatGPT (*W=0.975, p=0.389*). Table 1 shows the descriptive statistics of the scores split by group. In addition, variances were homogeneous for the final exam scores of the two groups, as assessed by Levene's test for equality of variance, F(1,62)=0.558, (p=0.458), as shown in Table 2.

Table 2. Homogeneity of Variances Test of the Final Exam Scores

F	df	df2	p
0.558	1.0	62	0.458

Since the data was normally distributed and the variances were homogenous, thus an independent t-test was applicable. Table 3 shows the result of the independent t-test. An independent t-test revealed a t-statistic of 1.13, with df=62 (p=0.264). This suggests that there is a difference between the groups' means. Specifically, the group that did not use ChatGPT had a higher mean test score (X=20.2, SD=4.62) than those that used ChatGPT (M=18.5, SD=5.28). However, it is relatively small, as Cohen's d of 0.333 indicates. Hence, the observed difference is not likely to have practical significance. Therefore, this study's results indicate no statistically significant difference between the two groups' mean test final exam scores. This analysis indicates that ChatGPT can help improve students' understanding of programming, which is aligned with Jin and Kim (2023) and Jing et al. (2024). Furthermore, the student's proficiency level in programming is at par with that of those who are more proficient.

Table 3. Independent Samples t-Test of the Final Exam Scores

t-statistics	df	p	Cohen's d
1.13	62.0	0.264	0.333

RQ2: What are the Students' Experiences Using ChatGPT to Learn to Program?

The second research inquiry explored students' experiences using ChatGPT to accomplish the programming tasks, providing valuable insights into their perspectives and engagements with the LMM. Twenty-six (55%) of the 49 students who used ChatGPT shared their experiences. Four themes emerged from the thematic analysis.

Appreciative and Satisfied

The students shared a range of experiences, with the majority expressing gratitude and contentment with ChatGPT's help in solving the different programming tasks, which parallels that of Tlili et al. (2023), Ngo (2023), and Husain (2024), who stated that ChatGPT is generally regarded positively by students. One student noted, "ChatGPT helped me figure out and understand the right codes for my given problem. It made the process of solving it easier." They were particularly pleased with its ability to provide detailed explanations of the programs. Another student remarked, "ChatGPT assisted me in coding and helped me understand how to code more." Similarly, a student mentioned, "It helped me understand how iterations work well." Another student shared, "My experience using ChatGPT to answer has been beneficial. ChatGPT helped me solve programming problems that I found confusing. I believe that it's a good guide to use because it helped me learn about how to execute the codes." Figure 4 shows an example of the interaction between the student and ChatGPT.

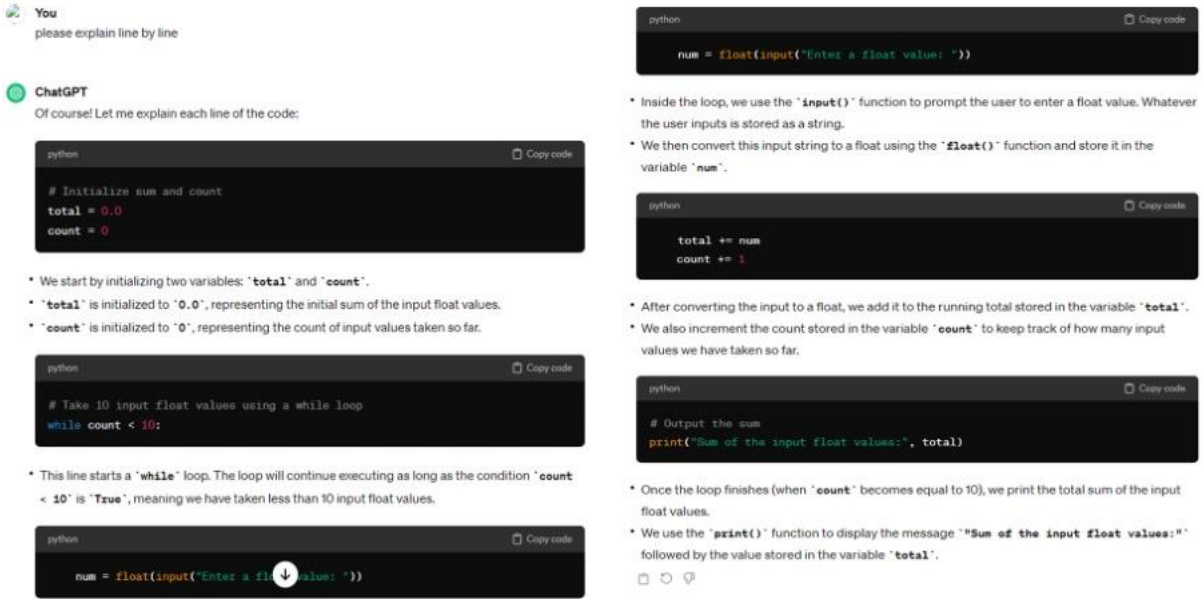


Figure 4. Example Interaction Between the Student and ChatGPT

Furthermore, students found ChatGPT a valuable resource when the teacher was unavailable. As one student mentioned, "[ChatGPT] was very helpful to me during the times when our teacher was unavailable or when he was handling other students [in] the class. The responses were immediate, allowing me to accomplish my task in time." Students generally agreed that ChatGPT has improved their learning experience by offering efficient solutions and enhancing their comprehension of programming principles.

The claim that ChatGPT helps improve students' understanding of programming principles is further supported by their experiences emphasizing ChatGPT's efficacy in delivering and generating Python programs, debugging, and explaining the lines of code. Students generally agreed that ChatGPT has improved their learning experience by offering efficient solutions and enhancing their comprehension of programming principles. Moreover, ChatGPT proved a great asset when the instructor was unavailable, enabling prompt task completion. As one student mentioned, "[ChatGPT] was very helpful to me during the times when our teacher was unavailable or

when he was handling other students [in] the class. The responses were immediate, allowing me to accomplish my task in time."

This finding is consistent with Ding et al. (2023), Jalil et al. (2023), Ngo (2023), Stojanov (2023), Wardat et al. (2023), and Zhu et al. (2023) that ChatGPT can be a tutor, assistant, and guiding tool to improve students' understanding of the lesson by providing rapid assistance to students in programming (Coello et al., 2024; Hartley et al., 2024; Jalil et al., 2023; Jin & Kim, 2023; Jing et al., 2024; Popovici, 2023).

Surprised and Amazed

The second theme revealed was that students who initially utilized ChatGPT were surprised by its capabilities, particularly its ability to understand and produce Python programs. The students were fascinated by the technology and impressed by its quick and correct generation of Python programs. One student remarked, "It is my first time using [ChatGPT] when coding and I am kind of surprised, because it gives the answer and it is kind of accurate. When eliminating the codes that shouldn't be used and you tell the AI to remove them, it does and the code works. It is very detailed and helpful, thus this would really help students who are having trouble in doing the task. I can say that ChatGPT is efficient." Another student shared, "I was fascinated to know that ChatGPT can provide an AI-written code that was acceptable with only a few corrections needed." This result was not reported in any existing literature on the use of ChatGPT.

Acknowledging Limitations

While students were surprised and amazed by ChatGPT's capabilities, specifically its fast response and generation of Python programs, they recognized the limitations of ChatGPT, including frequent inaccuracies or occasions where ChatGPT supplied irrelevant details. One student noted, "ChatGPT gives unnecessary details or overcomplicates the code when given instructions. Revision is almost always needed. ChatGPT also explains the code, but what is needed is just the output provided to the user." Another student shared, "I think it was super useful to use Chat GPT to both understand and solve a programming problem. I liked how it created a program when I told it to not use functions, commands, and boolean values, but only use while loops, conditional statements, and logical operators. But there are times that it is incorrect. Like in the code below, it included both num = 1 and end_condition = 1 which is technically the same (As said by my professor)." This outcome is congruent with the studies of Coello et al. (2024), Husain, 2024, Jalil et al. (2023), and Popovici (2023) in the field of programming.

There were occasions when there was a need for detailed directives or ChatGPT's responses deviated from their prompts, requiring students to give additional details and do follow-ups. Other students expressed similar sentiments, with one stating, "AI needs specific instructions to work as intended" and another mentioning, "It helps but it isn't 100% accurate, so I just use it as a guide." Figure 5 shows an example of this occurrence. These observations were substantiated by Coello et al. (2024), who argued that the effectiveness of ChatGPT is affected by the prompts provided by users, while Zhu et al. (2023) proposed that ChatGPT's incorrect responses can be

corrected by doing follow-up prompts.

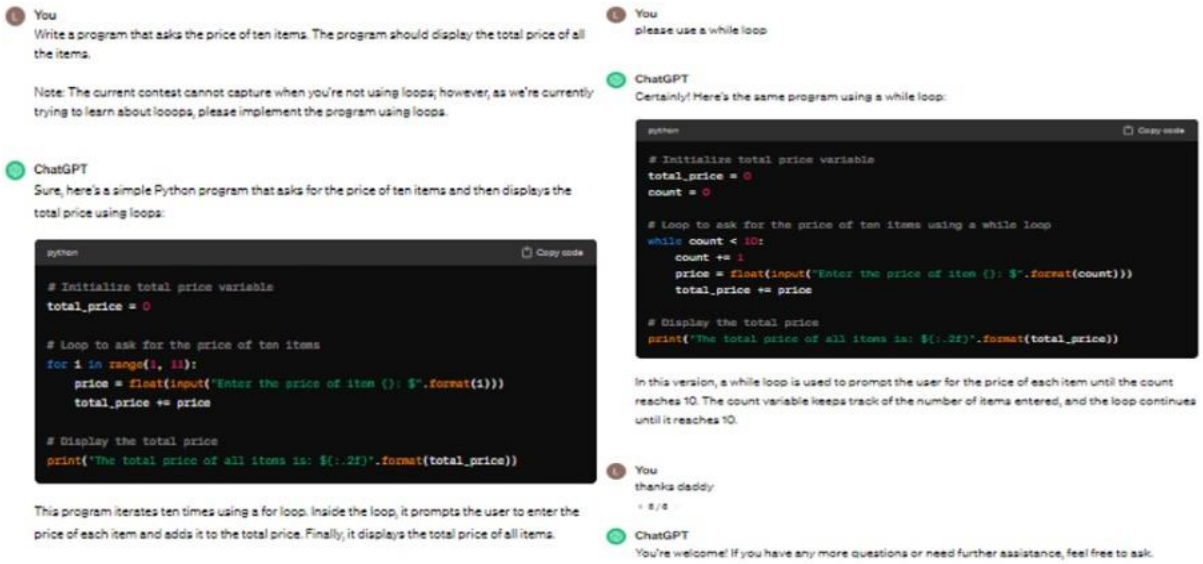


Figure 5. ChatGPT's Revising Its Response Based on the Student's Prompt

ChatGPT is a Person

Another theme uncovered was that some students treated ChatGPT as a person. This was reflected in the students' prompts. Students said, "Thank you," used the word "Please," expressed their emotions to ChatGPT, asked ChatGPT about its capabilities to generate codes according to the requirements, and called it "Daddy." Figure 10 shows example prompts of a student thanking and expressing his/her excitement for coding to ChatGPT. The students said, "ITZ PYTHON CODING TIME BABYYY YEAHHHHHH…!!! This result is consistent with the student's misconception reported by Ding et al. (2023) that the students see ChatGPT as a human or a robot, which can be attributed to the ChatGPT's human-like conversations and cognitive abilities.

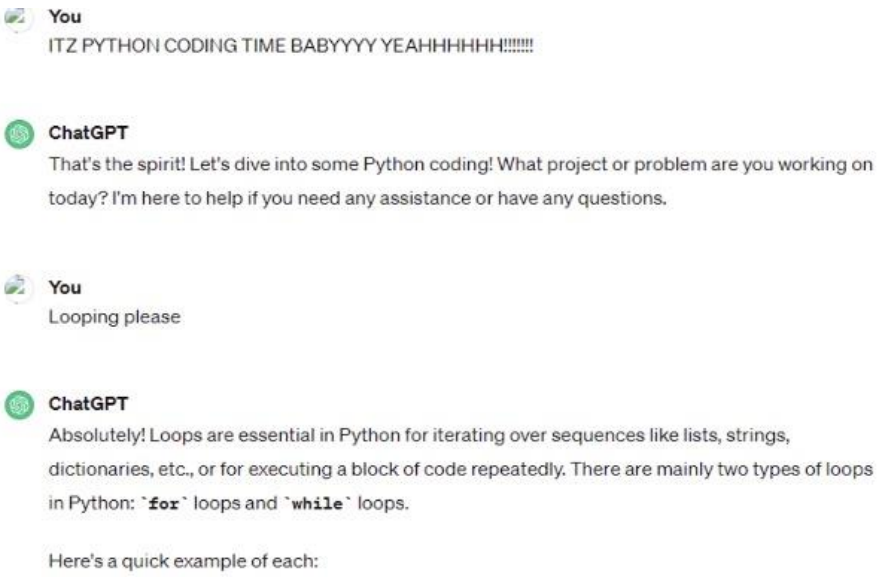


Figure 10. Sample Prompts of a Student Treating ChatGPT as a Person

Conclusion

The release of ChatGPT-3 in November 2022 by OpenAI attracted the attention of millions and made ChatGPT a headline. Researchers raised issues regarding the use of ChatGPT in the education sector. While there may be specific issues, it is widely recognized that ChatGPT has significant potential to improve the teaching and learning process. The main goal of this study is to find out how ChatGPT affects SHS students' learning of Python programming. The study looks at students' proficiency and experiences using ChatGPT as a teaching tool to shed light on the function of AI-driven tools in programming instruction. The study uncovered valuable results regarding ChatGPT's advantages and constraints. A considerable proportion of students employed ChatGPT for iterative statements, while its utilization for the term project was restricted due to the limitations set by the teacher. Although ChatGPT demonstrated promise, its efficacy relies on the user's prompts.

There was no significant difference in the final exam scores between students who utilized ChatGPT and those who did not, indicating the potential of ChatGPT to enhance comprehension of programming. The effectiveness of ChatGPT in providing solutions and improving understanding was demonstrated through students' experiences, which is consistent with earlier research. Although students recognized and valued ChatGPT's advantages, they also understood and highlighted its limits. Furthermore, students treated ChatGPT as a person due to its human-like responses. This research showed that education stakeholders can incorporate AI tools such as ChatGPT to enhance conventional teaching approaches in Python programming courses, offering supplementary assistance and direction to learners. However, it is essential to emphasize the importance of utilizing artificial intelligence technologies to develop critical thinking and independent problem-solving skills.

Recommendations

Although our research has provided valuable insights, it is crucial to recognize certain limitations. There were differences in the sizes of the groups, and just a small number of students participated. Moreover, future research could delve into the long-lasting effects of ChatGPT on students' programming efficiency and experiences. Furthermore, exploring methods to boost ChatGPT's compliance with given instructions and minimize inaccuracies will increase its usefulness in educational environments.

Acknowledgment

The authors want to acknowledge the DLSU-IS SHS administrators for allowing this study.

References

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Coello, C. E. A., Alimam, M. N., & Kouatly, R. (2024). Effectiveness of ChatGPT in coding: A Comparative analysis of popular marge language models. *Digital 2024, 4*, 114–125.

https://doi.org/10.3390/digital4010005

Creswell, J.W. & Creswell, J.D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Thousand Oaks, CA: Sage

Ding, L., Li, T., Jiang, S., & Gapud, A. (2023). Students' perceptions of using ChatGPT in a physics class as a virtual tutor. *International Journal of Educational Technology in Higher Education, 20*(1), 63. https://doi.org/10.1186/s41239-023-00434-1

Hartley, K., Hayak, M., & Ko, U. H. (2024). Artificial intelligence supporting independent student Learning: An evaluative case study of ChatGPT and learning to code. *Education Sciences, 14(2)*, 120. https://doi.org/10.3390/educsci14020120

Husain, A. (2024). Potentials of ChatGPT in computer programming: Insights from programming instructors. *Journal of Information Technology Education, 23*, Article 2. https://doi.org/10.28945/5240

Jalil, S., Rafi, S., LaToza, T. D., Moran, K., & Lam, W. (2023). ChatGPT and software testing education: Promises & perils. *2023 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, 4130–4137. https://doi.org/10.1109/ICSTW58534.2023.00078

Jauhiainen, J. S., & Guerra, A. G. (2023). Generative AI and ChatGPT in school children's education: Evidence from a school lesson. *Sustainability (Basel, Switzerland), 15*(18), 14025. https://doi.org/10.3390/su151814025

Jin, J., & Kim, M. (2023). GPT-empowered personalized elearning system for programming languages. *Applied Sciences, 13*(23), 12773. https://doi.org/10.3390/app132312773

Jing, Y., Wang, H., Chen, X., & Wang, C. (2024). What factors will affect the effectiveness of using ChatGPT to solve programming problems? A quasi-experimental study. *Humanities & Social Sciences Communications, 11*(1), 319. https://doi.org/10.1057/s41599-024-02751-w

Kiryakova, G., & Angelova, N. (2023). ChatGPT—A challenging tool for the university professors in their teaching practice. *Education Sciences, 13*(10), 1056. https://doi.org/10.3390/educsci13101056

Liang, Y., Zou, D., Xie, H., & Wang, F. L. (2023). Exploring the potential of using ChatGPT in physics education. *Smart Learning Environments, 10*(1), 52. https://doi.org/10.1186/s40561-023-00273-7

Ngo, T. T. A. (2023). The Perception by university students of the use of ChatGPT in education. *International Journal of Emerging Technologies in Learning, 18*(17), 4–19. https://doi.org/10.3991/ijet.v18i17.39019

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods, 16*(1), 1–13. https://doi.org/10.1177/1609406917733847

OpenAI (2022, November 30). *Introducing ChatGPT*. Retrieved April 12, 2024, from https://openai.com/blog/chatgpt

Pokkakillath, S., & Suleri, J. (2023). ChatGPT and its impact on education. *Research in Hospitality Management, 13*(1), 31–34. https://doi.org/10.1080/22243534.2023.223957

Popovici, M. D. (2023). ChatGPT in the classroom. Exploring its potential and limitations in a functional programming course. *International Journal of Human–Computer Interaction*, 1–12. https://doi.org/10.1080/10447318.2023.2269006

Stojanov, A. (2023). Learning with ChatGPT 3.5 as a more knowledgeable other: An autoethnographic study. *International Journal of Educational Technology in Higher Education, 20*(1), 35. https://doi.org/10.1186/s41239-023-00404-7

Teubner, T., Flath, C. M., Weinhardt, C., van der Aalst, W., & Hinz, O. (2023). Welcome to the era of ChatGPT et al: The prospects of large language models. *Business & Information Systems Engineering, 65*(2), 95–101. https://doi.org/10.1007/s12599-023-00795-x

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*(1), 15–24. https://doi.org/10.1186/s40561-023-00237-x

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Revised ed.). Harvard Univ Pr.

Wardat, Y., Tashtoush, M. A., AlAli, R., & Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(7), em2286-. https://doi.org/10.29333/ejmste/13272

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

Author Information

Joel B. Jalon, Jr.
https://orcid.org/0009-0005-2465-3764
De La Salle University - Integrated School
Philippines
Contact e-mail: *joel.jalon@dlsu.edu.ph*

Goodwin A Chua
https://orcid.org/0009-0009-7066-5685
De La Salle University - Integrated School
Philippines

Myrla de Luna Torres
https://orcid.org/0009-0007-1322-4184
De La Salle University - Integrated School
Philippines