

Enhancing Math Learning with AI: ChatGPT's Impact on Number Base Conversion Comprehension

Sireesha Prathi Gadapa

School of Mathematics, Actuarial and Quantitative Studies, Asia Pacific University Malaysia

Email: sireesha.prathi@apu.edu.my

Salwani Binti Mohd Daud

University Malaysia of Computer Science and Engineering, Malaysia

Email: salwani.daud@unimy.edu.my

Bryan Thien Chek Hui

School of Mathematics, Actuarial and Quantitative Studies, Asia Pacific University Malaysia

Email: bryan.thien@apu.edu.my

Morampudi Rama Tulasi Raju

Technical Director, Maven IT Solutions Malaysia

Email: mrtraju@mavenitsolutions.com.my

To Link this Article: <http://dx.doi.org/10.6007/IJARPED/v13-i3/21642>

DOI:10.6007/IJARPED/v13-i3/21642

Published Online: 18 June 2024

Abstract

This research investigates the effectiveness of using questions generated by Chat Generative Pre-trained Transformer (ChatGPT), a large language model, to improve student comprehension and performance in number base conversions. A mathematics course with 340 participants underwent pre-assessment to establish the baseline knowledge. Following a lecture, a subgroup of 170 students received a formative assessment with varying difficulty levels of ChatGPT-generated questions. Post-assessment was administered to all students, aiming to discern improvements and evaluate the impact of ChatGPT-generated questions on learning outcomes. The Wilcoxon-singed rank test revealed a significant enhancement in student performance after the intervention, while the Man-Whitney U test showed that the students who were exposed to the ChatGPT-generated questions scored better than those who did not, and its effectiveness was not gender-biased. More importantly, the findings show that the group exposed to ChatGPT-generated questions not only performed better but also demonstrated a deeper understanding of the material compared to their peers. This suggests that ChatGPT's questions effectively targeted diverse learning needs, potentially by providing personalized learning experiences within the assessment. This research contributes to Artificial Intelligence (AI) integration in education, highlighting ChatGPT's potential to create effective and insightful learning experiences in mathematics courses. This research

aims to equip students with essential skills and knowledge that are directly applicable to their academic and professional pursuits in computing and technology. The potential contributions of this research include enhancing students' computational proficiency, fostering innovative problem-solving abilities, and preparing them for the complex demands of the technology industry.

Keywords: ChatGPT, Student Performance, Wilcoxon-singed Rank Test, Man-Whitney U Test, Number Base Conversions.

Introduction

The integration of AI technologies in education holds immense potential to revolutionize teaching and learning practices. By offering personalized learning experiences and adaptive instruction, AI enables educators to address the diverse needs of students more effectively, ultimately fostering deeper understanding and mastery of academic material.

AI Technologies in Educational Settings

Artificial Intelligence (AI) has progressively permeated various facets of modern life, including education (Chen et al., 2020). Within educational settings, AI technologies offer innovative solutions to traditional challenges, revolutionizing teaching and learning methodologies (Luan et al., 2020). These technologies encompass a diverse array of applications, ranging from intelligent tutoring systems to natural language processing models like ChatGPT. Intelligent tutoring systems (ITS) represent one prominent example of AI in education (Alkhatlan & Kalita, 2018). These systems leverage AI algorithms to provide personalized instruction, adapting content and pacing to meet individual student needs (Akyuz, 2020). Through real-time assessment of student performance and comprehension, ITS can offer targeted interventions and scaffolding, fostering deeper understanding and retention of academic material (Rajaram, 2021). Furthermore, AI-powered educational platforms utilize machine learning algorithms to analyze vast datasets of student interactions and outcomes. By identifying patterns and trends within this data, these platforms can offer valuable insights to educators, enabling informed decision-making and instructional refinement. Additionally, AI technologies facilitate the development of adaptive learning environments, where instructional content dynamically adjusts based on individual learner characteristics and progress (Rane et al., 2023).

Importance of Personalized Learning and Adaptive Instruction

Personalized learning and adaptive instruction represent fundamental pillars of effective education in the 21st century (Natal'ya & Kopceva, 2023). Traditional one-size-fits-all instructional approaches often fail to accommodate the diverse learning styles, paces, and preferences of students (Haniya & Roberts-Lieb, 2017). As a result, some students may struggle to engage with the material, while others may become disengaged due to a lack of challenge. AI technologies address this challenge by enabling personalized learning experiences tailored to the unique needs of each student (Rane et al., 2023). By utilizing algorithms that analyze student data and behavior, AI systems can deliver customized content, pacing, and feedback. This personalized approach not only enhances student engagement and motivation but also promotes deeper learning and mastery of academic concepts. Moreover, adaptive instruction facilitated by AI technologies allows for real-time adjustments to instructional strategies based on student performance and comprehension (Gligorea et al., 2023). Through continuous assessment and feedback mechanisms, educators

can identify areas of difficulty and provide targeted interventions to support student progress. This adaptive approach ensures that students receive the support and resources they need precisely when they need them, maximizing learning outcomes.

Role of ChatGPT-generated Formative Assessment in Enhancing Learning Outcomes

The integration of ChatGPT-generated formative assessment holds significant implications for enhancing learning outcomes in several key ways. Firstly, the adaptive nature of AI-generated questions allows for personalized learning experiences, catering to the diverse needs and learning styles of individual students (Willis, 2023). By presenting questions of varying difficulty levels and content domains, ChatGPT facilitates differentiated instruction, ensuring that each student receives challenges commensurate with their abilities. ChatGPT-generated formative assessment questions, coupled with personalized feedback from educators, synergistically enhance learning outcomes. This collaborative interplay between AI technology and human expertise exemplifies the symbiotic relationship at the heart of modern education. While AI facilitates the generation of formative assessment questions that are adaptive, dynamic, and personalized, educators provide the essential human touch in interpreting results, providing contextually relevant feedback, and fostering deeper understanding.

ChatGPT as a Pedagogical Tool

ChatGPT represents a powerful pedagogy tool that empowers educators to create dynamic, personalized, and engaging learning experiences (Mohebi, 2024). Its ability to generate human-like text and engage in natural language interactions makes it uniquely suited for enhancing various aspects of teaching and learning (Adiguzel et al., 2023). As educators explore innovative approaches to instruction, ChatGPT offers a range of applications that contribute to student engagement, comprehension, and personalized learning experiences. By dynamically creating questions across different difficulty levels and content domains, ChatGPT facilitates differentiated instruction, allowing educators to assess student understanding and provide targeted feedback in real-time (Supriyadi & Kuncoro, 2023). This adaptive approach promotes active learning and metacognitive awareness, empowering students to take ownership of their learning journey.

Role of Number Base Conversions in Computing

Number base conversions play a fundamental role in computing by facilitating the representation and manipulation of data in various formats. In computing, binary (base-2), hexadecimal (base-16), and decimal (base-10) systems are commonly utilized for encoding and processing information. Converting between these number bases enables the interpretation of data stored in different formats, such as converting binary data to hexadecimal for readability or decimal for arithmetic operations. Additionally, number base conversions are integral to addressing memory management, bitwise operations, and encoding schemes in computer systems. Mastery of number base conversions is essential for understanding the underlying mechanisms of computing and software development.

Research Objectives

This research seeks to investigate how integrating ChatGPT-generated questions as pedagogical tools can enhance computing students' understanding of number base conversions. By examining pre and post-assessment outcomes, the study aims to assess the

impact of ChatGPT-generated questions on student learning outcomes and pedagogical practices in computing education.

1. To investigate the effectiveness of ChatGPT-generated assessments as a pedagogical tool on students' performance.
2. To determine the impact of ChatGPT-generated assessments as a pedagogical tool on students' performance.
3. To examine the impact of ChatGPT-generated assessments as a pedagogical tool on students' performance based on their gender.

Research Questions

1. Is there a significant difference in students' performance for those who used ChatGPT-generated assessments as a pedagogical tool?
2. How does the use of ChatGPT-generated assessment as a pedagogical tool impact the overall students' performance?
3. Is there a significant difference in the effectiveness of ChatGPT-generated assessment as a pedagogical tool between male and female students' performance?

Literature Review

AI technologies are revolutionizing education by enhancing teaching quality, assessing student learning, and improving educational value. The integration of Artificial Intelligence in Education (AIED) applications is rapidly advancing due to the development of computing technologies (S., A. 2023). AI tools like ChatGPT are being utilized to improve teaching quality, align student learning with societal needs, and enhance assessment quality, ultimately raising the overall quality of education (Liu et al., 1 C.E.) (Alneyadi & Wardat, 2023). Personalized learning through AI technologies enables tailored learning experiences, improved student outcomes, increased engagement, and enhanced educational experiences, emphasizing the importance of ethical considerations and responsible implementation for equitable access to quality education for all learners (Anekal et al., 2023). Implementing AI-supported activities in schools is crucial for transforming teaching and preparing students with digital competencies, requiring teachers to assess the appropriateness of AI tools for achieving educational goals.

The use of ChatGPT in formative assessments has shown positive impacts on students' overall test scores in summative assessments. The studies conducted by Alneyadi & Wardat et al (2023); Alneyadi & Wardat (2023); Sanchez-Ruiz et al (2023) both found that students who had access to ChatGPT demonstrated significant enhancements in their scores on post-tests measuring student achievement. Alneyadi et al (2023) reported that the experimental group, which utilized ChatGPT for assistance, displayed higher mean scores across all subscales compared to the control group. Similarly, Sanchez-Ruiz et al (2023) found that students quickly adopted ChatGPT as a valuable tool, exhibiting high confidence in their responses and general usage in the learning process. These findings suggest that the use of ChatGPT in formative assessments can positively impact students' overall test scores in summative assessments. ChatGPT-generated formative assessment plays a crucial role in enhancing learning outcomes in higher education. By utilizing ChatGPT's capabilities, educators can design effective multiple-choice questions that cater to formative assessment needs (Rasul et al., 2023). This form of assessment is particularly valuable in asynchronous online learning environments, where adult learners can benefit from ChatGPT's support in setting goals, locating resources, and monitoring progress toward self-directed learning. Additionally, the

integration of alternative assessments that consider the use of AI chatbots like ChatGPT can provide a more comprehensive understanding of student's knowledge and skills acquired in informal learning settings, ultimately contributing to improved learning outcomes in language programs. Despite concerns about academic integrity, when used ethically, ChatGPT can foster supportive learning environments, promoting deeper learning and better outcomes for students (Lin, 2023).

The integration of ChatGPT into education has had a positive impact on student engagement and learning outcomes. Students have found ChatGPT to be a valuable learning tool, providing assistance with course content and helping them understand difficult concepts (Alneyadi & Wardat, 2023). It has been particularly beneficial in disciplines such as social science, where extensive research and analysis are required [3]. ChatGPT has been shown to enhance student achievement, with the experimental group displaying significant improvements in their scores compared to the control group (Siregar et al., 2023). ChatGPT has shown promise as a pedagogical tool in mathematics, particularly in enhancing educational success (Wardat et al., 2023). It has demonstrated varying levels of accuracy in responding to mathematics questions, excelling in simpler topics but facing challenges with more complex concepts (Dao & Le, 2023). While ChatGPT's potential in education is acknowledged globally, caution is advised due to limitations such as inaccuracies and ethical concerns (Azaria et al., 2023). In engineering education, students have embraced ChatGPT, recognizing its usefulness but expressing concerns about its impact on developing lateral competencies essential for future engineers (Sánchez-Ruiz et al., 2023).

Overall, ChatGPT's integration into mathematics education offers opportunities for improved learning experiences, but careful consideration of its limitations and the need for strategic usage is essential to ensure its effective and responsible implementation.

Methodology

In this research, formative assessment questions were generated with the assistance of artificial intelligence (AI), specifically using ChatGPT. The AI-generated questions are tailored to target various difficulty levels and accommodate the diverse needs of the student cohort. These questions are designed to assess student understanding of number base conversions, a complex mathematical concept. By using ChatGPT to generate questions, educators can ensure that students are challenged appropriately and receive a range of questions that test different aspects of their knowledge.

Research Design

The research employs a quasi-experimental design to assess the effectiveness of integrating ChatGPT-generated assessments in improving students' understanding of number base conversions. Three assessment phases are conducted over a period of five weeks: pre-assessment, ChatGPT assessment, and post-assessment. **Figure 1** illustrates a quasi-experimental design focusing on the comparison between an experimental group and a control group.

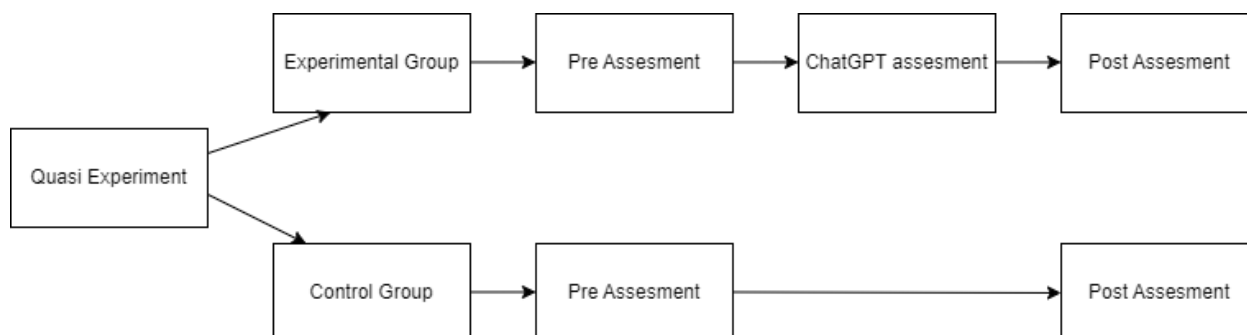


Figure 1

Quasi Experiment Design: In the quasi-experimental design depicted in **Figure 1**, two groups are compared: the experimental group and the control group. The experimental group undergoes a three-step process, beginning with a pre-assessment to establish their baseline understanding of number base conversions. Following this, they engage in a ChatGPT assessment designed to evaluate their comprehension and performance using dynamically generated questions. Finally, the experimental group completes a post-assessment to measure any changes in their understanding after the instructional intervention. Conversely, the control group follows a simpler trajectory, participating only in the pre-assessment at the outset to establish their baseline understanding. They then proceed directly to the post-assessment after a comparable period without exposure to the ChatGPT-generated questions or other instructional interventions. This design allows for a comparison between the experimental group, which receives the ChatGPT intervention, and the control group, which does not. By examining the differences in pre- and post-assessment scores between the two groups, educators evaluate the effectiveness of the ChatGPT intervention in enhancing student comprehension and performance on number base conversions."

Participants: The research comprised a total of 340 students enrolled in a Mathematics for Computing (MCFC) module at Asia Pacific University of Technology and Innovation (APU), Malaysia. These students were pursuing various BSc (Hons) computing and technology-related programs in Year 1, semester 2. The student cohort was divided into two groups: an experimental group and a control group, each consisting of 170 students. This division facilitated a comparative analysis of the effectiveness of ChatGPT-generated questions in enhancing students' understanding of number base conversions. To ensure unbiased distribution, random selection methods were employed to assign students to either the experimental or control group.

Experimental Group: The experimental group comprised 170 students who participated in the ChatGPT assessment. These students were exposed to ChatGPT-generated questions designed to assess their comprehension of number base conversions after receiving a lecture on the topic.

Control Group: The control group also consisted of 170 students who did not participate in the ChatGPT assessment. They underwent the same lecture on number base conversions as the experimental group but did not engage with ChatGPT-generated questions. The control group served as a baseline against which the effectiveness of ChatGPT-generated questions could be evaluated.

Learning Outcomes Description

Table 1 describes the learning outcomes of the "Number Base System" topic of the MCFC module, which is a foundational component of BSc (Hons) computing and technology-related programs in Year 1, semester 2. These outcomes are designed to equip students with essential

knowledge and skills in understanding and manipulating number bases, essential for their future studies and professional endeavors in computing and technology fields. Through this module, students will gain proficiency in conceptualizing and converting between decimal, binary, octal, and hexadecimal number systems, laying a solid foundation for advanced mathematical and computational principles in subsequent coursework and practical applications.

Table 1

Assessment Procedures

Learning outcomes	Description
Understand the Concepts of Number Bases	<ul style="list-style-type: none"> Define the principles underlying each number base system, including place value and positional notation. Describe the characteristics of decimal, octal, binary, and hexadecimal number systems.
Convert Numbers Between Different Number Bases	<ul style="list-style-type: none"> Convert numbers between decimal, binary, octal, and hexadecimal representations.

Assessment serves as a crucial component in evaluating students' comprehension and learning outcomes. In this research, a comprehensive assessment strategy was devised to measure students' understanding of number base conversions before and after exposure to instructional interventions. The assessment procedures encompassed three main phases: pre-assessment, ChatGPT assessment, and post-assessment.

Pre-Assessment: At the onset of the study, all 340 students underwent a pre-assessment to establish a baseline understanding of number base conversions. The pre-assessment comprised a series of questions covering various aspects of the topic, designed to gauge students' initial proficiency levels and identify areas for improvement.

ChatGPT Assessment: Subsequently, a subset of 170 students was randomly selected from the total cohort to participate in the ChatGPT assessment. This assessment occurred after students received a lecture on number base conversions. ChatGPT, utilizing its natural language processing capabilities, generated questions of varying difficulty levels tailored to individual student needs. This interactive assessment aimed to provide personalized learning experiences and offer insights into students' comprehension levels.

ChatGPT-Generated Questions

Incorporating Bloom's Taxonomy Bloom (2014) into the generation of questions for research can significantly enhance the cognitive rigor and depth of learning outcomes. According to the Malaysian Qualifications Agency (MQA) standards, questions categorized under different levels of Bloom's Taxonomy correspond to specific cognitive processes. For instance, questions prompting students to "find" information typically align with the C1 level, representing basic recall or comprehension of factual knowledge. Conversely, questions prompting students to "convert" information, requiring them to apply, analyze, or evaluate concepts, align with the C2 level, reflecting higher-order thinking skills. By utilizing ChatGPT to generate questions aligned with Bloom's Taxonomy, researchers can ensure that the assessment tasks effectively target diverse cognitive domains and foster deep conceptual understanding.

Figure 2 depicts the interaction between the educator and ChatGPT during the formative assessment. The educator presents a question prompt related to binary to decimal base conversion to ChatGPT.

ChatGPT 3.5 ▾



You

Please give one easy, one medium, and one hard-level question on binary to decimal base conversion. For easy level use the keyword find, medium level convert, high level convert. Allocate 1 mark for Easy level question, 2 marks for both medium and hard level questions.

Figure 2

Figure 3 illustrates the response generated by ChatGPT in response to the educator's question prompt.

ChatGPT 3.5 ▾

1. A file requires 256 bytes of storage. Find the binary equivalent of the decimal base number.
(1 mark)

Medium Level:

2. The representation of a network route is 110001101.0010110110. Convert the binary base number into Hexadecimal.
(2 marks)

Hard Level:

3. A security token is CEED.E9. Convert the Hexadecimal base number into decimal.
(2 marks)

Figure 3

Table 2

Describes the ChatGPT-generated question to cover various aspects of number base conversions, including binary, hexadecimal, and decimal systems. Questions were categorized into easy, medium, and high-level difficulty levels to accommodate the diverse needs of students

Easy	Medium	Difficult
A file requires 256 bytes of storage. Find the binary equivalent of the decimal base number. (1 mark)	The representation of a network route is 110001101.0010110110. Convert the binary base number into Hexadecimal. (1 mark)	A security token is CEED.E9. Convert the Hexadecimal base number into decimal. (2 marks)
A program is loaded into memory at address 64. Find the octal equivalent of the decimal base number. (1 mark)	The configuration setting for a software application is 356.42. Convert the octal base number into binary. (1 mark)	A scientific calculator displays the number 3145.125. Convert the octal base number into decimal. (2 marks)

You are assigned to meeting room 226. Find the Hexadecimal equivalent of the decimal base number. (1 mark)	A measurement device displays its reading as 456.25. Convert the octal base number into decimal. (1 mark)	The version number of the software is 57321.176. Convert the octal base number into Hexadecimal. (2 marks)
You are transmitting data 1101101 over a network. Find the decimal equivalent of the binary base number. (1 mark)	In a communication system, control signals are represented as 4721.35. Convert the octal base number into Hexadecimal. (1 mark)	In a computer program, a program counter is set to BAD. ADD. Convert the Hexadecimal base number into an octal. (2 marks)
You have a binary file with the representation 110101. Find the octal equivalent of the binary base number. (1 mark)	A hardware device has a configuration code as B2C1.3D. Convert the Hexadecimal base number into binary. (1 mark)	A code used in a communication system is 11001011100101.0110010111001010. Convert the binary base number into Hexadecimal. (2 marks)

Post-Assessment

Following the completion of the lecture series, all 340 students underwent a post-assessment to measure their learning outcomes. The post-assessment, akin in format to the pre-assessment, covered the same content areas of number base conversions. By comparing pre and post-assessment results, the effectiveness of instructional interventions, including the ChatGPT assessment, could be evaluated.

Data Collection

Data collection occurs throughout the five weeks, with pre-assessment conducted in Week 1, ChatGPT assessment in Week 3, and post-assessment in Week 5. Assessment data includes students' responses to multiple-choice, short-answer, and problem-solving questions, as well as demographic information such as gender, nationality, and admission qualifications.

Data Analysis

Quantitative analysis involves comparing pre- and post-assessment scores to assess changes in students' understanding of number base conversions. Additionally, the effectiveness of ChatGPT-generated questions is evaluated by comparing the performance of students who participated in the ChatGPT assessment with those who did not. Descriptive statistics, including means, standard deviations, counts, and minimum and maximum values are used to summarize the assessment results. Before the implementation of statistical methods on the data to answer the research questions, Shapiro-Wilk test is conducted to check the normality of the data distribution. If the data is normally distributed, parametric tests, such as paired-sample t-test and independent t-test will be adopted to examine differences between groups and assess the significance of findings. However, if the data is not normally distributed, the non-parametric tests such as the Wilcoxon signed-rank test and Mann-Whitney U test will be deployed instead.

By implementing this methodology, this research aimed to evaluate the impact of ChatGPT-generated questions on students' learning outcomes and pedagogical practices in the context of computing education.

Results

In this section, the results are presented in the following manner, where the first section shows the descriptive analysis of the dataset, followed by its normality testing to determine the suitable statistical analyses to be used. Finally, various statistical analyses will be conducted on the data to address the research questions.

Descriptive Analysis

This dataset consists of 202 males and 138 females, constituting a total of 340 participants. Based on **Table 3**, females appeared to have a slightly higher score than males for both pre- and post-assessments. It is also worth noting that the minimum score for post-assessment has increased to 5 for males and 7 for females respectively.

	N	Pre-Assessment Score				Post-Assessment Score			
		Mean	Std. Deviation	Min	Max	Mean	Std. Deviation	Min	Max
Male	202	13.78	5.27	0	20	16.32	3.54	5	20
Female	138	15.25	4.63	0	20	16.89	2.97	7	20

Table 3

Table 4 shows the descriptive statistics of the pre- and post-assessment scores for each control and experimental group. The control and experimental groups show comparable results in the pre-assessment, whereas the experimental group appeared to score better than the control group in the post-assessment.

Table 4

Effectiveness of the ChatGPT-generated assessment on students' performance (RO1)

	N	Pre-Assessment Score				Post-Assessment Score			
		Mean	Std. Deviation	Min	Max	Mean	Std. Deviation	Min	Max
Control	170	14.35	5.10	0	20	14.71	3.12	5	20
Experimental	170	14.64	5.01	0	20	18.39	2.39	8	20

The first crucial step before the selection of statistical techniques is to assess the normality of data distribution. **Figure 4** and **Figure 5** illustrate the distribution of pre- and post-assessment scores for the control and experimental groups respectively. Evidently, the data does not follow a normal distribution. To further validate our findings, the Shapiro-Wilk test is computed on the data. Table 5 shows the results of the normality test. Since all the p-values are less than 0.05, this indicates that the data exhibits a significant deviation from normality. Hence, non-parametric tests are implemented on the data for further analyses.

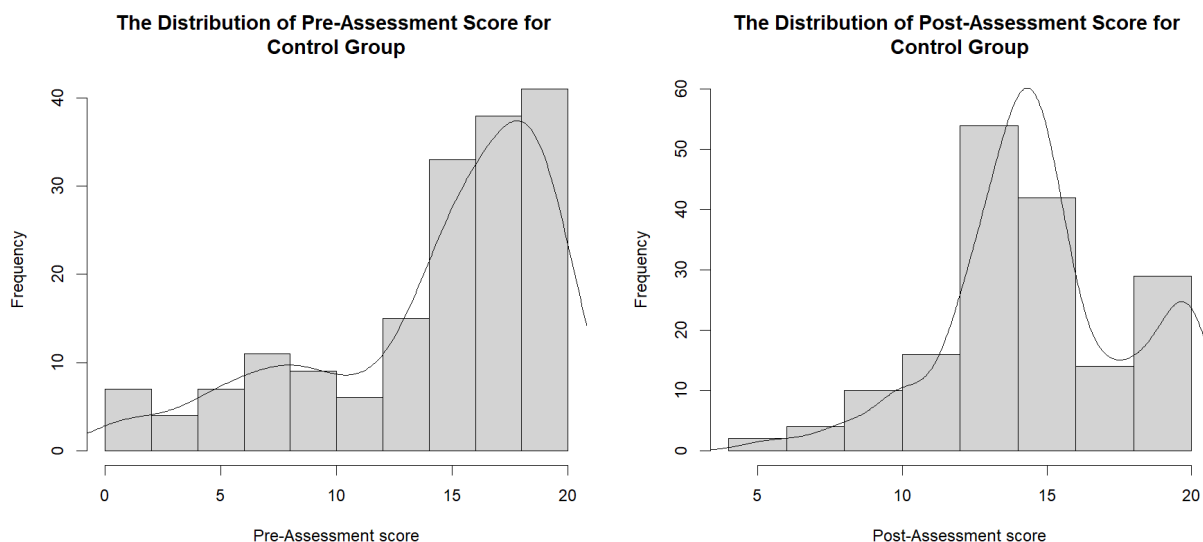


Figure 4

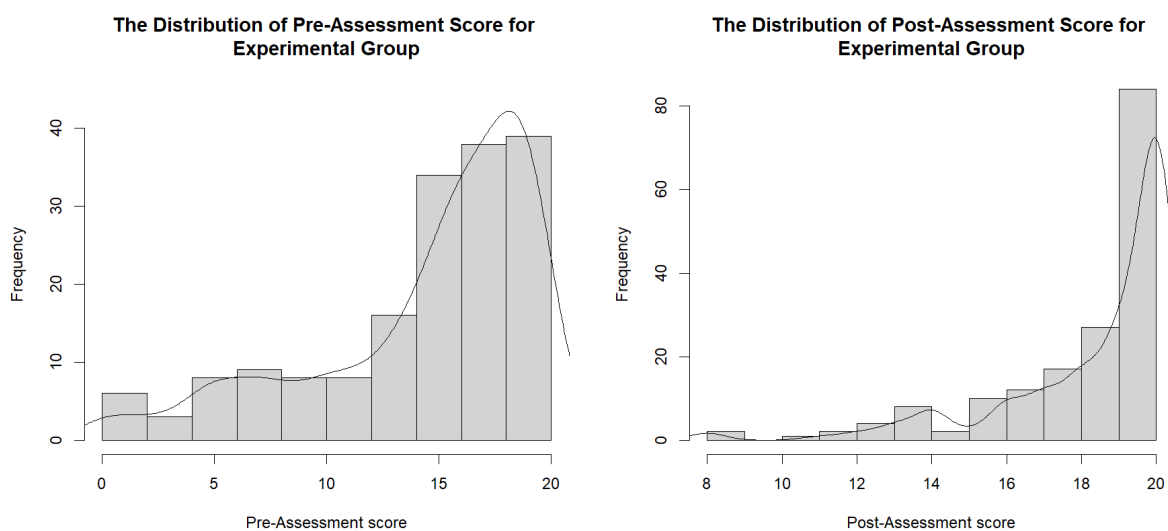


Figure 5

Table 5

Data	Shapiro-Wilk Statistics	p-Value	Normal Distribution
Pre-Assessment score for Control Group	0.85708	1.245e-11	No
Post-Assessment score for Control Group	0.94575	4.041e-06	No
Pre-Assessment score for Experimental Group	0.84694	5.121e-12	No
Post-Assessment score for Experimental Group	0.7173	< 2.2e-16	No

The Wilcoxon signed-rank test will be used to answer the first research question: “Is there a significant difference in students’ performance for those who used ChatGPT-

generated assessments as a pedagogical tool?" This test will compare the pre-assessment and post-assessment scores within each group to determine if there is a significant impact of ChatGPT-generated assessments on students' performance. Table 6 shows the results of the Wilcoxon signed-rank test for each group.

Table 6

Comparison	Test Statistics	p-Value	Significant Difference
Pre- and Post-Assessment Score of Control Group	6304	0.8171	No
Pre- and Post-Assessment Score of Experimental Group	1106	< 2.2e-16	Yes

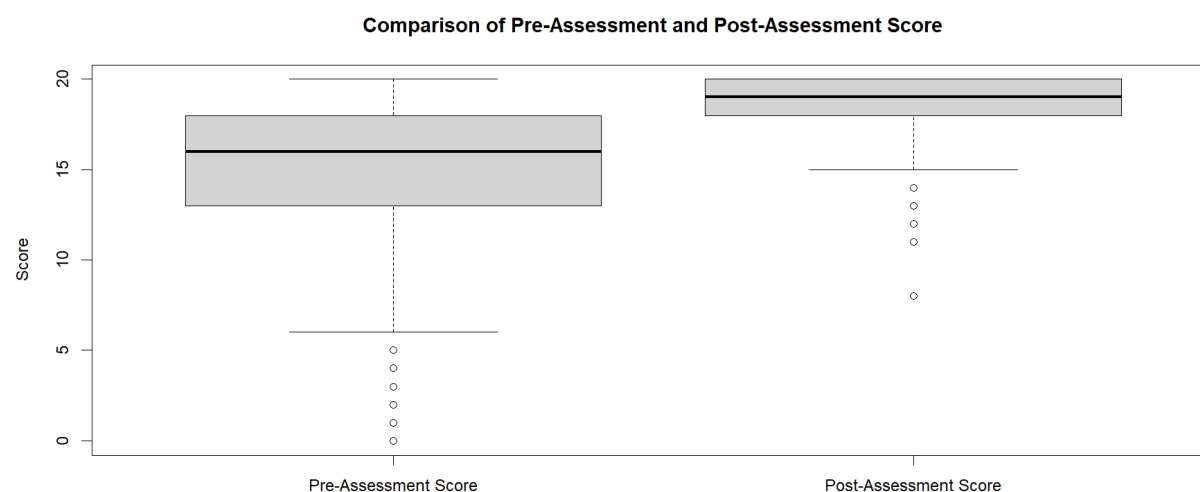


Figure 6

According to the results shown in Table 6, there is a significant difference between the pre- and post-assessment scores of the experimental group students who used ChatGPT-generated assessments as a pedagogical tool since the p-value computed is less than 0.05. On the other hand, it shows no significant difference for the control group since the p-value is greater than 0.05. Since the average post-assessment score is higher than the pre-assessment score for the experimental group as shown in Figure 6, this implies that the implementation of ChatGPT-generated assessments is effective in improving students' performance.

Impact of the ChatGPT-generated assessment on students' performance (RO2)

The post-assessment score of both control and experimental groups will be used for the second research question: "How does the use of ChatGPT-generated assessment as a pedagogical tool impact the overall students' performance?" Due to the non-normal distribution of the data, as evidenced in Figures 4 and 5 and confirmed by Table 5, the Mann-Whitney U test is adopted to test the significant difference in the post-assessment performance between both control and experimental groups to answer the second research question.

Table 7

Comparison	Test Statistics	p-Value	Significant Difference
Post-Assessment Score between Control and Experimental Groups	23654	< 2.2e-16	Yes

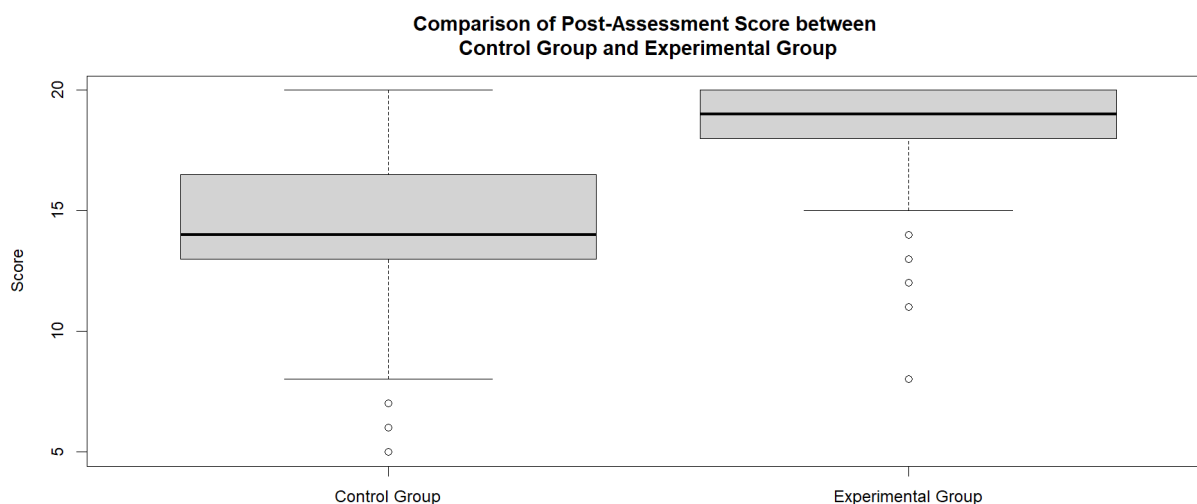


Figure 7

Based on the results shown in **Table 7**, there is a significant difference in the post-assessment score between both control and experimental groups as the p-value generated is less than 0.05. Since the average post-assessment score of the experimental group is higher than that of the control group as depicted in **Figure 7**, the results indicate that the students who used the ChatGPT-generated assessment performed better than those who did not.

Effectiveness of the ChatGPT-generated assessment on gender (RO3)

The post-assessment score of both male and female students in the experimental group will be utilized for the last research question: “Is there a significant difference in the effectiveness of ChatGPT-generated assessment as a pedagogical tool between male and female students’ performance?” Typically, it is essential to perform normality testing to determine the most appropriate analysis method to be used on the data. Figure 8 shows the distribution of post-assessment score for both males and females in the experimental group.

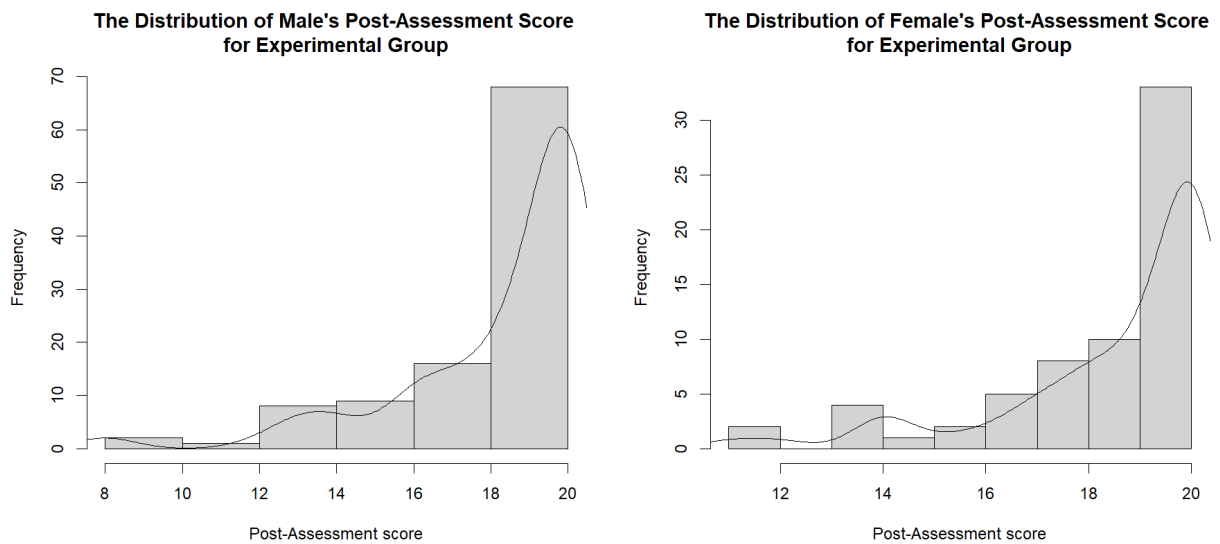


Figure 8

Table 8

Data	Shapiro-Wilk Statistics	p-Value	Normal Distribution
Post-Assessment score of Male for Experimental Group	0.71433	6.265e-13	No
Post-Assessment score of Female for Experimental Group	0.72358	9.444e-10	No

Apparently, the data does not adhere to a normal distribution and this statement is further supported by the Shapiro-Wilk test since the p-value generated is less than 0.05 as shown in Table 8. Since the data fails the normality testing, the Man-Whitney U test is adopted to address the last research question.

Table 9

Comparison	Test Statistics	p-Value	Significant Difference
Post-Assessment Score between Males and Females of Experimental Group	3271	0.7072	No

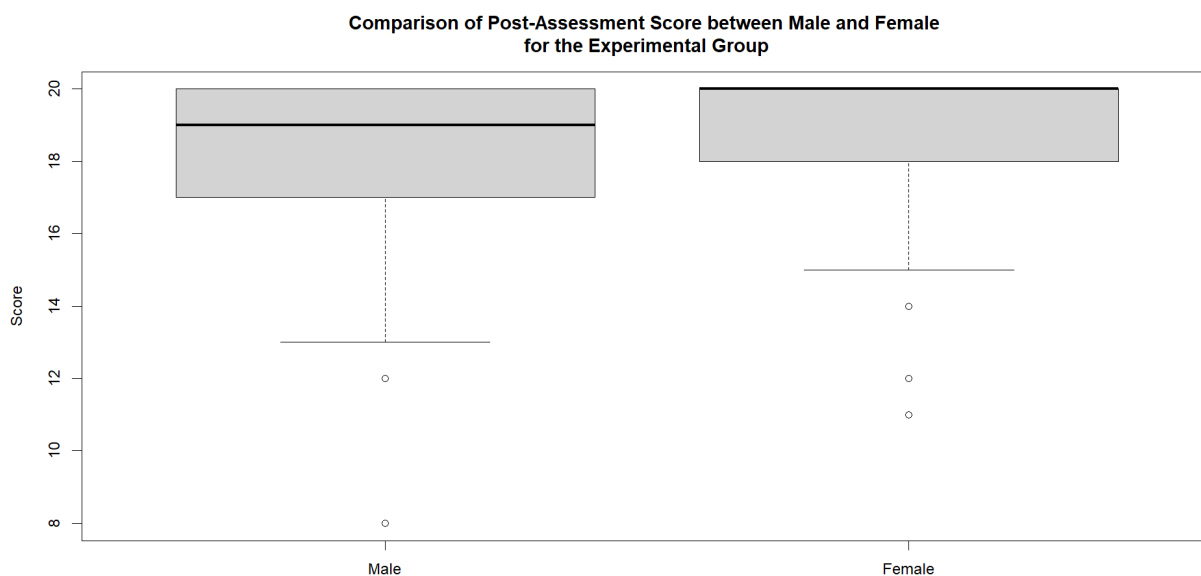


Figure 9

According to the results shown in **Table 9**, there is no significant difference in the post-assessment score between the male and female students of the experimental group since the p-value computed is more than 0.05, which implies that the ChatGPT-generated assessment is effective for all students in getting good academic results regardless of their gender. This phenomenon is illustrated in Figure 9, where the post-assessment scores for both male and female students of the experimental group are comparable.

Conclusion

Given the dynamic and ever-evolving nature of technology, it is imperative to explore innovative approaches to enhance teaching and learning experiences in computing disciplines. ChatGPT, as an advanced natural language processing tool, presents a unique opportunity to transform traditional assessment practices by offering personalized, interactive, and adaptive learning experiences. This research sought to assess whether integrating ChatGPT-generated questions into pedagogical practices could lead to improved student engagement, comprehension, and retention of mathematical concepts, such as number base conversions. Furthermore, by analyzing the effectiveness of ChatGPT-generated questions, this research aimed to inform pedagogical strategies and curriculum design in computing education. The study aims to equip students with essential skills and knowledge that are directly applicable to their academic and professional pursuits in computing and technology. The potential contributions of this study include enhancing students' computational proficiency, fostering innovative problem-solving abilities, and preparing them for the complex demands of the technology industry. Understanding how AI-driven technologies like ChatGPT can complement traditional teaching methods is essential for preparing students for the demands of the digital age and ensuring their success in future careers in computing and technology-related fields. By leveraging the capabilities of ChatGPT, educators can create a more engaging, effective, and adaptable learning environment that meets the needs of today's students and equips them with the competencies required to excel in an increasingly digital world. This research underscores the importance of integrating AI tools into education to not only enhance learning outcomes but also to prepare students for the challenges and opportunities of the future.

References

- Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology, 15*(3), ep429. <https://doi.org/10.30935/cedtech/13152>
- Akyuz, Y. (2020). Effects of Intelligent Tutoring Systems (ITS) on Personalized Learning (PL). *Creative Education, 11*(6), Article 6. <https://doi.org/10.4236/ce.2020.116069>
- Alkhatlan, A., & Kalita, J. (2018). *Intelligent Tutoring Systems: A Comprehensive Historical Survey with Recent Developments* (arXiv:1812.09628). arXiv. <https://doi.org/10.48550/arXiv.1812.09628>
- Alneyadi, S., & Wardat, Y. (2023). ChatGPT: Revolutionizing student achievement in the electronic magnetism unit for eleventh-grade students in Emirates schools. *Contemporary Educational Technology, 15*(4), ep448. <https://doi.org/10.30935/cedtech/13417>
- Azaria, A., Azoulay, R., & Reches, S. (2023). *ChatGPT is a Remarkable Tool—For Experts* (arXiv:2306.03102). arXiv. <https://doi.org/10.48550/arXiv.2306.03102>
- Can ChatGPT Improve Creative Problem-Solving Performance in University Students?* (2023). <https://doi.org/10.31234/osf.io/9z2tc>
- Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. *IEEE Access, 8*, 75264–75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
- Dao, X.-Q., & Le, N.-B. (2023). *Investigating the Effectiveness of ChatGPT in Mathematical Reasoning and Problem Solving: Evidence from the Vietnamese National High School Graduation Examination* (arXiv:2306.06331). arXiv. <https://doi.org/10.48550/arXiv.2306.06331>
- Gligorea, I., Cioca, M., Oancea, R., Gorski, A.-T., Gorski, H., & Tudorache, P. (2023). Adaptive Learning Using Artificial Intelligence in e-Learning: A Literature Review. *Education Sciences, 13*(12), Article 12. <https://doi.org/10.3390/educsci13121216>
- Haniya, S., & Roberts-Lieb, S. (2017). Differentiated Learning: Diversity Dimensions of e-Learning. In *E-Learning Ecologies*. Routledge.
- Hanum Siregar, F., Hasmayni, B., & Lubis, A. H. (2023). The Analysis of Chat GPT Usage Impact on Learning Motivation among Scout Students. *International Journal of Research and Review, 10*(7), 632–638. <https://doi.org/10.52403/ijrr.20230774>
- Li, J., Gui, L., Zhou, Y., West, D., Aloisi, C., & He, Y. (2023). *Distilling ChatGPT for Explainable Automated Student Answer Assessment* (arXiv:2305.12962). arXiv. <https://doi.org/10.48550/arXiv.2305.12962>
- Lin, X. (2023). Exploring the Role of ChatGPT as a Facilitator for Motivating Self-Directed Learning Among Adult Learners. *Adult Learning, 10451595231184928*. <https://doi.org/10.1177/10451595231184928>
- Liu, Y., Baucham, M., Liu, Y., & Baucham, M. (1 C.E., January 1). *AI Technology: Key to Successful Assessment* (ai-technology) [Chapter]. <https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-6684-8292-6.ch016>; IGI Global. <https://www.igi-global.com/gateway/chapter/www.igi-global.com/gateway/chapter/323556>
- Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S. J. H., Ogata, H., Baltes, J., Guerra, R., Li, P., & Tsai, C.-C. (2020). Challenges and Future Directions of Big Data and Artificial Intelligence in Education. *Frontiers in Psychology, 11*. <https://doi.org/10.3389/fpsyg.2020.580820>
- Mohebi, L. (2024). Empowering learners with ChatGPT: Insights from a systematic literature exploration. *Discover Education, 3*(1), 36. <https://doi.org/10.1007/s44217-024-00120-y>

- Natal'ya, Kopceva. (2023). Effective learning strategies in the 21st century: The concept of A.A. Ozdogru. 2023, 2(2), 28–55. <https://doi.org/10.31804/2782-540x-2023-2-2-28-55>
- On ChatGPT: what promise remains for multiple choice assessment? | *Journal of Learning Development in Higher Education*. (n.d.). Retrieved March 1, 2024, from <https://journal.aldinhe.ac.uk/index.php/jldhe/article/view/1009>
- Rajaram, K. (2021). Learning Interventions: Collaborative Learning, Critical Thinking and Assessing Participation Real-Time. In K. Rajaram (Ed.), *Evidence-Based Teaching for the 21st Century Classroom and Beyond: Innovation-Driven Learning Strategies* (pp. 77–120). Springer. https://doi.org/10.1007/978-981-33-6804-0_3
- Rane, N., Choudhary, S., & Rane, J. (2023). *Education 4.0 and 5.0: Integrating Artificial Intelligence (AI) for Personalized and Adaptive Learning* (SSRN Scholarly Paper 4638365). <https://doi.org/10.2139/ssrn.4638365>
- Rasul, T., Nair, S., Kalendra, D., Robin, M., Santini, F. de O., Ladeira, W. J., Sun, M., Day, I., Rather, R. A., & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning and Teaching*, 6(1), Article 1. <https://doi.org/10.37074/jalt.2023.6.1.29>
- Sánchez-Ruiz, L. M., Moll-López, S., Nuñez-Pérez, A., Moraño-Fernández, J. A., & Vega-Fleitas, E. (2023). ChatGPT Challenges Blended Learning Methodologies in Engineering Education: A Case Study in Mathematics. *Applied Sciences*, 13(10), Article 10. <https://doi.org/10.3390/app13106039>
- Supriyadi, E., & Kuncoro, K. S. (2023). Exploring the future of mathematics teaching: Insight with ChatGPT. *Union: Jurnal Ilmiah Pendidikan Matematika*, 11(2), Article 2. <https://doi.org/10.30738/union.v11i2.14898>
- The Influence of ChatGPT on Social Science Students: Insights Drawn from Undergraduate Students in the United States. (2023). *Indonesian Journal of Innovation and Applied Sciences*. <https://doi.org/10.47540/ijias.v3i2.878>
- 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>
- Willis, V. (2023). The Role of Artificial Intelligence (AI) in Personalizing Online Learning. *Journal of Online and Distance Learning*, 3(1), Article 1. <https://doi.org/10.47941/jodl.1689>