AFRICAN JOURNAL OF THEORY AND PRACTICE OF EDUCATIONAL RESEARCH (AJTPER)
Print ISSN: 2630-6565
e-Copy ISSN: 2992-3638
VOL. 13 JUNE, 2025

# The Use of AI to Support Computational Thinking in Mathematical Tasks Using GeoGebra in Measurement and Evaluation Course

**Mary Patrıck Uko (Ph.D)[1], Dr Mfon Edet Thompson[2] & Dr Glory Udo Evans[3]**

Department of Educational Foundation, School of Education, College of Education, Afaha Nsit, Akwa Ibom State, Nigeria
E-mail; maryuko@akscoe.edu.ng ORCID 0009-0000-6314-4376
+2348028502943
[2]E-mail;Mfonthompson84@gmail.com +2348022048155
[3]E-mail;evansgloryu@gmail.com +2348035678520

**Abstract**
This study explored the potential of artificial intelligence (AI), specifically ChatGPT, to improve computational thinking (CT) skills in mathematics education using GeoGebra. It examined how AI can integrate CT into core statistical concepts in a measurement and evaluation course in a college of education in Akwa Ibom State, Nigeria. Using an exploratory mixed-methods research design, to gain insight of the subject matter since it is not well defined, 100 second-year undergraduate students from the School of Science, with an emphasis on those with a background in computer science and access to the necessary technological resources, from a population of 2,674, were purposively sampled. To ensure the proper representation of the sample, participants were drawn from diverse academic backgrounds like mathematics, physics, biology giving credence to prior experience with GeoGebra, and familiarity with AI tools to contextualize the findings. To enhance the comprehension of the functionalities of this design, the researchers employed an educational design research (EDR) approach. This methodology allowed for a detailed exploration of participant experiences and the integration of quantitative and qualitative data. Quantitative data on CT performance and qualitative insights on AI interactions were analysed using a descriptive qualitative method and chi-square. Results indicate that ChatGPT facilitated the creation of GeoGebra commands, promoted algorithmic thinking, debugging, and independent problem solving. However, challenges such as command errors and inconsistent responses from AI have been observed. The study concludes that AI, combined with GeoGebra, promises differentiated and personalized lessons, highlighting its role in improving CT in mathematics education. Future research should explore its broader applications and address technical challenges for better integration.
**Keywords:** ChatGPT, Computational Thinking, Educational Design Research, GeoGebra, Mathematics

## Introduction

The integration of computational thinking (CT) in education has become a crucial aspect of modern teaching, particularly in STEM disciplines. CT provides students with essential problem-solving skills, allowing them to decompose problems, recognise patterns, and devise algorithmic solutions. In mathematics education, software tools such as GeoGebra have been instrumental in developing these skills by allowing students to visualise and interact with mathematical concepts dynamically (Ziatdinov & Valles, 2022). Recent advances in artificial intelligence (AI), particularly conversational models such as ChatGPT, have also expanded the potential for personalised and adaptive learning experiences. AI tools can act as virtual tutors, helping students build, correct, and understand mathematical models with software such as GeoGebra. For example, ChatGPT can guide students in generating GeoGebra commands step by step, improve correction processes, and adapt to individual learning needs (Cascella, Montomoli, Bellini, & Bignami, 2023; Yunianto et al., 2023).

In mathematics education, AI models such as Chat Generative Pre-Trained Transformer (AI-ChatGPT) can significantly improve computational thinking (CT) by helping students build, correct, and understand GeoGebra commands. GeoGebra is a dynamic software tool that enables students to create visual models for complex mathematical concepts. AI can act as an interactive tutor to guide students through this process, developing critical thinking (CT) skills such as algorithmic thinking, problem decomposition, and debugging. AI-ChatGPT can help students develop their algorithmic thinking by guiding them through the step-by-step construction of GeoGebra instructions. For example, a student learning to model a quadratic function or a pie chart must know how to plot points, set parameters, and create visual curves. AI-ChatGPT can provide step-by-step instructions on how to enter these commands, as well as explanations of why each command is needed, which promotes understanding and comprehension of algorithms (Castelvecchi, 2023; Udofia & Uko, 2018; Yunianto et al., 2024). Troubleshooting and error detection are among the most crucial functions of AI-ChatGPT. It can analyse syntax or logic errors in a GeoGebra application and suggest corrections. For example, if a student enters a rule or command incorrectly, AI-ChatGPT can identify the error and guide how to correct it. This can improve CT performance by helping students understand not only how to correct it but also why it happened, developing problem-solving skills that are essential to CT and the subject (Rane, 2023).

Problem analysis, debugging, error identification, and decomposition are areas where AI-powered tools like ChatGPT can assist. For complex mathematical problems, AI-ChatGPT can help students break down problems into smaller, more manageable pieces. For example, when working on creating a geometric model, AI-ChatGPT can guide students through breaking down the task into steps, such as identifying axes, selecting points, and drawing each line or shape. This allows students to approach the problem systematically, thereby strengthening CT skills through an organised incremental progression (Cotton, Cotton, P. & Shipway, 2023; Wardat et al., 2023).

Adaptive learning and personalisation are another aspect of AI-ChatGPT. Its interaction is adaptive, meaning it can adapt to a student's individual learning needs, space, and level of understanding. This adaptive quality allows the AI to provide personalised help, addressing specific areas where a student may struggle with CT concepts or GeoGebra commands. This flexibility fosters differentiated learning, helping students develop their CT skills more effectively by responding to their unique challenges (Shabunina et al., 2023). Acting as a responsive and interactive tutor in GeoGebra, AI-ChatGPT enhances students' computational thinking (CT) skills through continuous feedback, enabling them to visualise and solve mathematical problems independently.

In mathematics education, AI assists students in breaking down complex problems into manageable tasks. This systematic approach strengthens CT skills and fosters critical thinking and creativity. The integration of CT with AI such as ChatGPT and GeoGebra can significantly enhance students' understanding of statistical concepts and data evaluation in courses focused on educational measurement. Decomposition is the process of breaking down a complex problem into more straightforward, manageable steps. In mathematics, it involves developing a process for calculating statistical measures or geometric transformations that can be implemented in software such as GeoGebra (Shute et al., 2017). For example, solving a complex equation may involve breaking it down into smaller equations, identifying and simplifying the terms (Ziatdinov & Valles, 2022).

In the classroom, CT fosters critical thinking, creativity, and autonomy by transitioning students from rote memorisation to exploratory learning, where they can utilise tools like GeoGebra to visualise mathematical concepts, model scenarios, and interact with their solutions dynamically in real-time. Research has shown that integrating CT into mathematics instruction can significantly improve student engagement and understanding by encouraging them to "think like mathematicians" who not only solve problems but also understand the processes behind them (Shute et al., 2017; Wadat et al., 2023). In the context of educational measurement and evaluation courses, integrating computational thinking (CT) with AI tools like ChatGPT and software such as GeoGebra can enhance student engagement and deepen their understanding of statistical concepts, critical measurements, and data evaluation techniques. Tools like GeoGebra enable students to visualise mathematical concepts dynamically. Recent advancements in artificial intelligence (AI), especially tools like ChatGPT, enhance personalised learning by acting as virtual tutors that help students with GeoGebra commands (Oliveares et al., 2024; Yunianto et al., 2023). This promotes critical thinking (CT) by encouraging students to approach data analysis systematically, understanding the role of each component in the larger model (Sohail et al., 2023).

A key part of computational thinking (CT) in education goes beyond simple calculations to interpreting results within academic contexts. Tools like GeoGebra help students visualise and explore mathematical concepts interactively (Wardat et al., 2023; Ziatdinov & Valles, 2022), improving their understanding and reducing cognitive load (Sohail et al., 2023). AI tools, such as ChatGPT, can assist students with step-by-step guidance in model building and error troubleshooting, fostering skills in algorithmic thinking and debugging. Recent studies by Dos Santos and Cury (2023), Sohail et al. (2023) and Wardat et al. (2023), highlight the benefits of using AI, especially in conjunction with GeoGebra for complex math tasks, as it offers real-time feedback that helps students learn independently (Botanas, Recio & Velez, 2024; Oliveares et al., 2024). However, challenges arise in aligning AI assistance with specific mathematical details, requiring careful guidance from educators. The current studies by Van Borkulo et al. (2021) and Yunianto et al. (2023) lack strong theoretical frameworks, which limits their broader application. Integrating frameworks like Seymour Papert's Constructivism and Shute et al.'s CT models can deepen understanding of AI's role in enhancing CT. Papert's framework emphasises hands-on learning through creating models, while Shute's CT framework outlines key components like abstraction, algorithmic thinking, and debugging that are practised when using GeoGebra. This study aims to fill existing gaps by applying these frameworks to assess how ChatGPT enhances CT skills during statistical tasks involving GeoGebra in a Measurement and Evaluation course. Overall, the collaboration between AI tools and educational resources promotes active learning and deeper engagement with complex mathematical concepts (Adel Ashan & Davidson, 2024; Ali, 2024). However, challenges exist in ensuring AI assistance aligns with specific mathematical nuances, requiring educators to carefully facilitate its use (Baidoo-Anu & Owusu Ansah, 2023; Li et al., 2023; Wollny et al., 2021; Yunianto et al., 2024). This highlights the need for teachers to facilitate the use of AI with care, ensuring that students accurately interpret AI assistance in their educational context.

**Visual Representation of the Frameworks**

The diagram below is a conceptual representation illustrating how Papert's constructionism and Shute et al.'s CT frameworks can be utilised to foster the integration of computational thinking (CT) with AI tools, such as ChatGPT and GeoGebra.

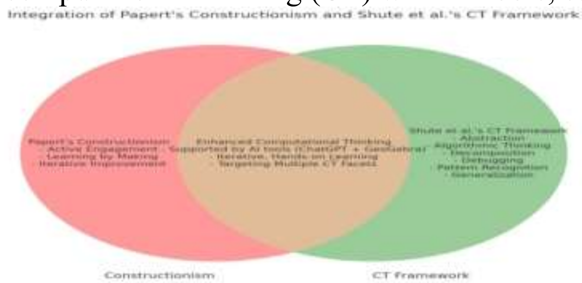


Figure 1. A conceptual diagram showing the overlap between Papert's structure and Shute et al.'s CT framework

**Illustration of the framework**

1. Central circle: represents the common goal of improving CT through active learning and practice. 2. Left panel (document structure): illustrates active participation, learning by doing and continuous improvement as key principles. Right Panel (Schott et al.

Aspects of CT): Shows the six aspects of CT (extraction, algorithmic reasoning, analysis, debugging, pattern recognition, and generalisation). 4. AI Tools (ChatGPT + GeoGebra): These tools act as an enabler, providing support and scaffolding for the frameworks. However, achieving this requires careful implementation to account for AI limitations and to foster meaningful student engagement in computational tasks. In summary, combining AI-ChatGPT with GeoGebra commands in educational measurement and evaluation courses enables students to gain hands-on experience in statistical modelling while developing essential computational thinking (CT) skills. By examining how AI tools can support CT skills in the context of GeoGebra-based tasks, the study aims to shed light on practical educational applications of AI and ways to enrich mathematics curricula. Specifically, the study aims to investigate how participants communicate with AI-ChatGPT during the Math CT activity, how AI-ChatGPT assists in problem-solving, the challenges encountered during the interaction, and the specific aspects of CT that are enhanced through engagement with AI. Four research questions and one hypothesis have been addressed in this study:

RQ1. How do participants communicate with AI-ChatGPT during the Math CT activity?

RQ2. How does AI-ChatGPT assist the participants in solving the Math CT activity?

RQ3.What are the potential challenges participants encounter when attempting to solve Math CT problems using AI-ChatGPT?

RQ4. What CT aspects are supported by interacting with AI-ChatGPT while solving the Math CT task?

HO1. There is no significant relationship between participants' groups(PG1, PG2, PG3, PG4, PG5) and their success or failure rates in completing GeoGebra tasks using AI-ChatGPT.

**Methodology**

This study stems from the curiosity of one of the researchers, who wondered how to help students who have developed a phobia of anything related to calculation. After careful consideration of these and the input received from the literature, conferences concerning the usefulness of AI-ChatGPT in education and her Masters' degree work on Geogebra, the researchers have now proceeded to explore the potential of incorporating AI-ChatGPT an emerging technology in generative artificial intelligence (AI), into teaching and learning of the course. This study examines the incorporation of computational thinking (CT) in the discipline of Measurement and Evaluation, specifically in Mathematics education. The objective is to contribute to the development of Math CT lessons that incorporate mathematics software, specifically GeoGebra and spreadsheets. A mixed-methods exploratory research design was employed to investigate the potential of AI-ChatGPT to enhance computational thinking (CT) in GeoGebra mathematics tasks. The educational design research (EDR) approach, as described by McKenney and Reeves (2018), was adopted to ensure the continuous refinement of the learning model, explore participants' experiences, and integrate quantitative and qualitative data. This study targeted 2,674 second-year undergraduate students enrolled in the EDU223

Measurement and Evaluation course in the College of Education, Akwa Ibom State, Nigeria. A purposive sampling technique was used to select 100 students from the school of Science, prioritising those with computer science combinations and access to the necessary technological resources. To ensure the representativeness of the sample, Participants were selected from different scientific fields such as mathematics, physics, and biology. Basic demographic data (age, previous experience with GeoGebra, and familiarity with AI tools) were collected to contextualise the findings. The reason for using a sample size of 100 participants in this study is that it provides a target group familiar with computational tools, such as GeoGebra and computers, which aligns with the study objectives.

The collected data were analysed quantitatively and qualitatively. Quantitative data focused on task completion rates, measured as the percentage of participants who completed the assigned tasks in GeoGebra. Debugging iterations: Records the number of errors during the first and last command attempts. Moreover, GeoGebra command success rates: Descriptive and frequency statistics were used to analyse trends across participant groups. Qualitative data were collected in the following ways: 1. Screen Recordings: Participants recorded their interactions with ChatGPT and GeoGebra to document problem-solving processes and adjustments. 2. Interviews: Structured interviews explored participants' perceptions of the effectiveness and challenges of ChatGPT. 3. AI-ChatGPT interaction logs: Conversations with immediate feedback were analysed to understand the role of different types of information in task success. The data were analysed quantitatively and qualitatively. The quantitative analysis used descriptive statistics (e.g., frequency, percentage) to analyse task accuracy, execution speed, and error detection repeatability. Chi-square tests were used to examine associations between participant groups (e.g., GeoGebra experience level) and success rates. The qualitative analysis employed thematic coding, where interaction logs and interview transcripts were coded to identify themes related to algorithmic thinking, problem-solving, abstraction, and pattern recognition. A triangulation of on-screen data, interaction logs, and interviews was also conducted to ensure the validity and reliability of the findings.

**Procedure**

1. Orientation and training: Participants attended a workshop on basic GeoGebra and ChatGPT and were assigned a standardised entry level. 2. Task: Participants were tasked with creating a pie chart using GeoGebra, given specific parameters (e.g., centre at (0,0), radius 5, and defined segments). The applications had to be specifically designed to work with ChatGPT to issue commands. 3. AI-ChatGPT Interaction: Participants were encouraged to experiment with different modifications to correct and modify GeoGebra applications while writing questions and answers. 4. Screen recording and reporting: Participants recorded their activities via screen recordings and submitted reports for analysis. 5. post-interview: focused on challenges, effectiveness of ChatGPT, and ideas for improvement. Ethical approval was obtained from the Institutional Research Ethics

Committee. Informed consent was obtained from all participants, and pseudonyms were used to ensure confidentiality.

**Results**

The results were reorganized into five discovery groups based on participant interaction, AI-ChatGPT use, and computational thinking (CT) aspects. Tables and graphs were added to summarise the results visually. The use of inferential statistics (chi-square test) was intended to strengthen the analysis, uncovering relationships between participant groups, communication methods, and success rates. Qualitative insights were examined through detailed excerpts from participants' interactions with ChatGPT, which illustrated key points, such as effective and ineffective requests and their impact on AI-generated ChatGPT responses. This added depth to the qualitative analysis, providing clear links between participants' strategies and outcomes.

**RQ1. First Prompt**

To answer research question one, the researchers investigated the participants' communication with AI-ChatGPT on their first prompt, as initial prompting plays a crucial role in informing AI-ChatGPT and prompting it to perform a task. The first prompt that the participants in Group 1 input to AI-ChatGPT was presented to determine whether their initial prompts were effective or not by evaluating them against an effective prompt. Individuals exhibit varying approaches to initiating conversations, as observed in this case, when the participants prompted AI-ChatGPT. Figure 2 represents the initial prompt provided by Participant in group 1 (PG1).

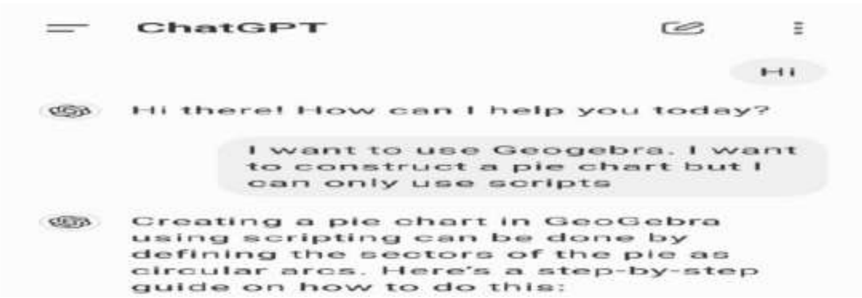


**Figure 2. Participant group 1 first prompt**

PG1 initiated the conversation with AI-ChatGPT by greeting it with 'Hi' or Hello and expressing their intention to utilise the GeoGebra. They did not provide any instructions regarding AI-ChatGPT's expertise in utilizing GeoGebra commands. Subsequently, they requested AI-ChatGPT to create a pie chart using scripts (Figure 3).


Figure 3

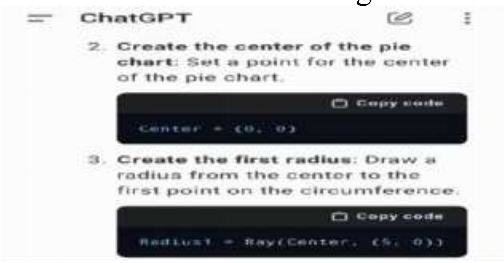



**Figure 4 & 5; How AI-ChatGPT responded to the use of script to construct the objects.**

Conversely, Participant group 2 (PG2) approached the task in a distinct manner. They directly requested assistance from AI-ChatGPT (Figure 4 & 5) but did not explicitly instruct it to assume the role of a GeoGebra expert. The AI-ChatGPT promptly answered and provided codes for this problem. The codes in PG1 and PG2 exhibit variances. In PG1, the initial value was 'center=(0,0)', but in PG2, it was 'C=Circle[(0,0),r]'. Based on the codes created by PG1 and PG2, it seems that AI-ChatGPT would produce different responses depending on the instructions.

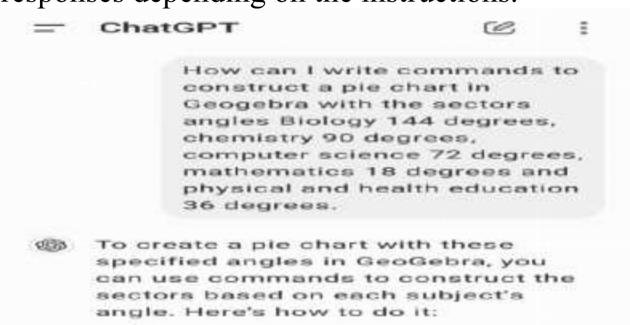

**Figure 6. Participant Group 2 first prompt**

To verify the validity of this assumption, the researchers proceeded with the remaining initial prompts and generated codes. Participants in Group 3 (PG3) utilised AI-ChatGPT Pro, a subscription-based platform, to request commands on how to generate a circle with a centre located at the coordinate (0,0) and a radius of 4. Figure 6 illustrates that AI-ChatGPT presented different codes for constructing a circle, namely 'Circle[(0,0),4]', which bear a resemblance to the codes of PG2's 'C=Circle[(0,0,r]', with the substitution of "r" with a numerical value.

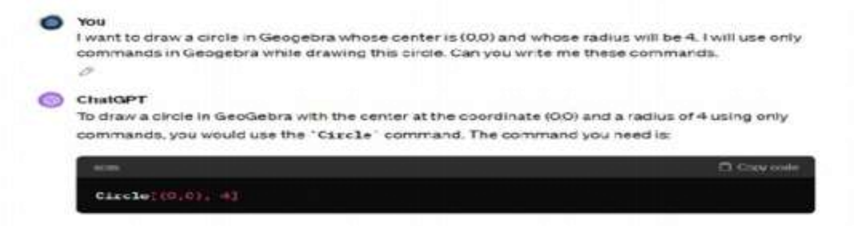


**Figure 7. Participant Group 3 first prompt**

In contrast, Participant Group 4 (P4) has a mathematics educational background similar to that of Participant Group 2 (PG2) and possesses basic knowledge of using GeoGebra. PG4 directly requested AI-ChatGPT to assist them in constructing objects without explicitly asking it to act as a GeoGebra expert. Figure 7 illustrates that the codes produced by AI-ChatGPT differ from those of P1, P2, and P3. The variable 'var radius=4' and other variables are present in P4's code.


**Figure 7. Participants' Group 4 first prompt**

Participant group 5 (PG5) has significant experience as mathematics teachers and demonstrates expertise in GeoGebra. They may have had prior knowledge of constructing the requested objects using GeoGebra. The researchers encouraged them to pretend to lack knowledge about the construction of the pie chart, instead, seek assistance from AI-ChatGPT. Figure 7 displays the initial prompt that they entered into AI-ChatGPT.


**Figure 8. Participants' group 5 first prompt**

Interestingly, PG5 instructed the AI-ChatGPT to possess expertise in GeoGebra and requested that it provide them with a sequence of GeoGebra commands for constructing the desired objects. The commands differ from prior participants due to the specified point A as A= (0,0). In the subsequent section, the researchers look deeper into the functioning of the codes generated by AI-ChatGPT.

**RQ2:** To answer the second research question, the researchers analysed the generated codes to see whether AI-ChatGPT provided codes.

It directly produced successful codes or not, and whether or not the codes differ from each participant. The prompts and communication methods used with AI-ChatGPT have an impact on both the responses and the generated codes. In contrast, PG2 generated distinct final codes. The sequential generation of codes to construct different components of the requested objects was displayed. Given the constraints of limited pages, the researchers have provided a summary of whether the final codes differ from those of all five participant groups. Table 1 presents the data regarding the final codes assigned to the participants, along with the number of iterations they underwent to obtain their final codes. The number of iterations indicates that the code has undergone some processes to arrive at a successful code that could construct the intended objects.

**RQ3:** To address the third research question, the researchers explored the challenges faced by participants in solving Math CT problems using AI-ChatGPT. They examined the success or failure of participants' final codes and the reasons for any issues that arose. The prompts and communication methods with AI-ChatGPT influenced the responses and codes generated. In one case in PG1, a command failed due to a small error in capitalisation. Another group, PG2, had success until a variable name conflict caused one of their codes to fail. In the PG4 group, an error with a specific command led to errors and failed constructions. However, participants in the PG5 group had good skills with GeoGebra and were able to argue effectively with AI-ChatGPT, even retraining it to correct errors. Each group's experience highlighted different aspects of working with AI tools.

RQ4: To address the fourth research question, we investigated the computational thinking (CT) facets that AI-ChatGPT supports using the CT framework by Shute et al. (2017). The researchers observed that AI-ChatGPT generated codes (GeoGebra commands) which need to be input into GeoGebra to construct the pie chart. The participants had to type in or copy and paste the code one by one into GeoGebra's input box. PG1, who had never used GeoGebra before, could learn how the GeoGebra commands worked as they created mathematical objects or representations. In this case, AI-ChatGPT can support the CT aspect 'algorithms' by Shute et al. (2017), which includes algorithm design to create a series of ordered steps to solve a problem. However, PG1's lack of basic understanding of GeoGebra commands constrained them from executing the proper commands. Despite this problem, PG1 learned how to program on GeoGebra. Another CT aspect, debugging, seems to be more prevalent in this study. As generated codes from AI-ChatGPT did not always work to construct the requested objects, the participants attempted to identify the problems and rectify them.

**Thematic Findings**

1. **Algorithmic thinking:** Participants demonstrated algorithmic thinking by constructing sequences of GeoGebra commands. For example, one participant noted: *"I understood that the sequence of commands had to follow a strict order: define the centre, define the radius, and then specify. The ChatGPT tips helped me think through the steps."* **2. Debugging:** This was a prominent theme, as many participants encountered errors in the AI-generated commands. One participant said: *"When ChatGPT suggested 'Sector (P, Q, n)', GeoGebra showed an error in ChatGPT for an alternative command. This process helped me understand the meaning of the mistakes."* **3. Abstraction:** Advanced participants, such as PG5, demonstrated abstraction by simplifying tasks. For example, instead of recreating the commands for each sector, they used a single variable to represent all angles: *"I grouped similar commands into a reusable formula and applied it to all sectors, which saved time."* **4. Pattern Recognition:** Some participants found that a repeated command structure was more efficient. One participant remarked: *"After the first few trials, I realised that the syntax for defining a point or a circle was similar. This pattern made it easier to create other shapes."* Insights Gained from Qualitative Analysis were: 1. Participants who had clear and structured questions were able to

receive detailed instructions from ChatGPT. 2. Limited experience with GeoGebra prevented some participants from debugging effectively, emphasising the need for foundational training. 3. Interaction styles (e.g., direct questions vs. exploratory prompts) influenced the depth of AI responses, highlighting the importance of prompt engineering.

**Quantitative Analysis**

1. **Participant Interactions and Communication Styles**

Participants employed various communication strategies with AI-ChatGPT, which influenced the quality of responses and outcomes. The result is presented in taqble 4 below.

**Table 4. Summary of the communication styles observed**

| Communication styles | Frequency in percentage | Examples |
|---|---|---|
| Direct question-based approach | 40% | How do i construct a circle with a radious of 5 in GeoGebra? |
| Explorative or iterative approach | 35% | Let's try this commands again could it work with different variables? |
| Instruction-based approach | 20% | You are a GeoGebra expert. Generate commands to plot a pie-chart with these parameters |
| Others / uncategorised | 5% | Informal prompts such as "I'm not sure how to start, can you help? |

Table 4 shows Participant Interaction and Communication Styles. Participants employed various communication strategies with AI-ChatGPT, which impacted the quality of responses and results. The success rate in creating correct GeoGebra commands among the participating groups varied significantly. Although PG5s had the highest success due to their prior knowledge of GeoGebra, PG1s struggled with the task, requiring many repetitions.

2. The data presented in Table 5 shows the prompts and communication methods that participants used with AI-ChatGPT.

**Table 5. The prompts and communication methods participants used with AI-ChatGPT**

| Communication methods | Frequency in percentage | Examples |
|---|---|---|
| Providing explanations and guidance | 50% | Clarifying math concepts or CT. |
| Suggesting step-by-step solutions | 30% | Providing detailed problem-solving steps. |
| Encouraging problem-solving strategies. | 15% | Recommending approaches for CT. |

| | | |
|---|---|---|
| Other forms of assistance. | 5% | Asking ChatGPT to possess expertise in GeoGebra and generate commands. |

The study found that well-structured and precise prompts were associated with the accuracy and usefulness of AI-ChatGPT responses. Participants such as PG5, who explained the work and function of the AI, received more precise GeoGebra instructions than PG1, who relied on imprecise instructions.

**3. Participants' debugging process and success rates**

Table 6 below shows a summary of the debugging process and success rates by participants.

**Table 6. Summary of the debugging process and success rates:**

| Group | Iterations to success | Success rate | Key challenges |
|---|---|---|---|
| PG1 | 16 | 0% | Misinterpreted prompts, AI provided invalid commands. |
| PG2 | 9 | 70% | Variable duplication; incomplete prompts. |
| PG3 | 4 | 100% | Strong command adaptation and abstraction. |
| PG4 | 6 | 60% | Incorrect usage of ‘point command. |
| PG5 | 5 | 100% | Advance debugging and prompt refinement. |

The study examines different levels of computational thinking (CT) use among the groups. PG5 stood out for modifying instructions, while PG1 and PG4 often helped fix bugs and improved AI responses. Pattern Extraction and Identification: PG3 excelled at identifying functional code patterns. The fourth research question examined the CT skills that AI-ChatGPT supported, based on the framework by Shute et al. (2017). It found that 35% of the participants used Problem decomposition, 30% engaged in pattern recognition, 25% practised Algorithmic thinking and 10% utilised Abstraction. The findings were visually represented in Figure 9, comparing the CT results among the groups.

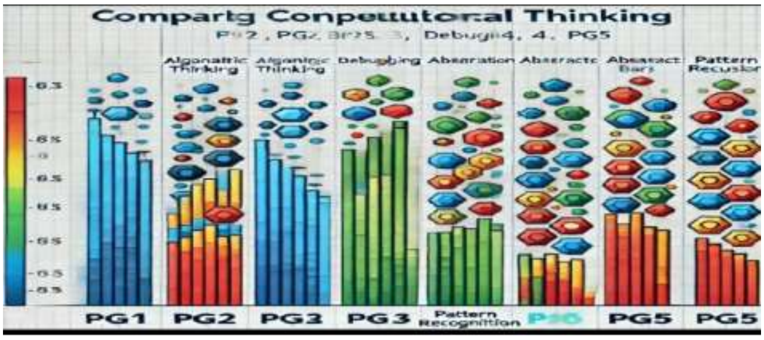


**Figure 9 Comparison of CT results between groups.**

**Explanation of the figure**

Figure 9 X-axis (participant group): Five participant groups (PG1, PG2, PG3, PG4, PG5) are shown to illustrate the learning pattern. 2. Y-axis (CT score): Scores out of 10 reflect performance in four areas of computational thinking: Algorithmic Thinking (blue), which measures participants' ability to generate step-by-step solutions. Debugging(red): Indicates participants' ability to identify and correct errors. Abstraction (green): Demonstrates the clarity of problems down to their most essential parts. Pattern Recognition (yellow): Indicates participants' ability to identify repetitive structures.

Key Findings: PG1 scores were very low in most areas, especially in debugging and abstraction, indicating a low level of experience and engagement with AI-ChatGPT. PG5 performs well across the board, with high scores in algorithmic reasoning and problem-solving, indicating strong prior knowledge and effective use of AI. PG2, PG3, and PG4 show mixed performance, with PG3 showing strong abstraction skills.

**HO1:** Table 7 below shows a contingency table summarising observed and expected frequencies of iterations for each group and success/failure outcomes, and chi-square ($\chi^2$) analysis used to test the association between observed and expected frequencies of iterations for each group and their success/failure outcomes.

**Table 7. A contingency table of Observed and Expected frequencies for each group and success/failure outcomes.**

| Group | **observed** Successful **frequencies** | Expected Successful Commands | **Observed** unsuccessful **frequencies** | Expected Unsuccessful commands |
|---|---|---|---|---|
| PG1 | 0 | 6.08 | 16 | 9.92 |
| PG2 | 6 | 5.70 | 9 | 9.30 |
| PG3 | 4 | 1.52 | 0 | 2.48 |
| PG4 | 4 | 3.80 | 6 | 6.20 |
| PG5 | 5 | 1.90 | 0 | 3.10 |

Table 8:Chi-Square Analysis for Success and Failure Rates Across Participant Groups.

| | Participant Groups |
|---|---|

| Rates | PG1 | PG2 | PG3 | PG4 | PG5 | N | df | $\chi^2$-cal | $\chi^2$-tab | Sig | p-value |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Observed (Expected) | | | | | | | | | | |
| Success Rate | 0(6.08) | 6(5.70) | 4(1.52) | 4(3.80) | 5(1.90) | 50 | 4 | 24.53 | 9.49 | 0.05 | .001 |
| Failure Rate | 16(9.92) | 9(9.30) | 0(2.48) | 6( 6.20) | 0(3.10) | 50 | | | | | |

**{ $\chi^2$ = 24.53, df = 4, p = 0.05>0.001} Note: PG1-PG5= Participant Groups.**

The study found significant differences in success rates among participant groups. PG1 and PG2 had the highest failure rates due to their lack of experience with GeoGebra and poorly structured prompts. PG1 had a 0% success rate, highlighting the challenges faced by individuals with limited prior knowledge. In contrast, PG3 and PG5 achieved 100% success rates, indicating that prior knowledge and clear prompts helped them use AI-ChatGPT effectively. PG5 also showed strong skills in debugging and abstraction. A chi-square test ($\chi^2$ = 24.53, p < 0.001) revealed a significant association between group success and rates, indicating that the differences were not due to chance. This highlights the importance of participant expertise and prompt quality. The findings suggest that training in prompt creation and GeoGebra skills is essential for effectively using AI tools. Future research should focus on providing tailored support for low-performing groups, such as PG1, and investigate how personalised AI feedback can help bridge knowledge gaps.

The significant difference between the observed and expected values, particularly for PG3 and PG5, reinforces the conclusion that success rates vary significantly across groups. This highlights how prior knowledge (e.g., PG5) or lack thereof (e.g., PG1) affects the effective use of AI-ChatGPT for GeoGebra tasks. Findings in Percentages 1. Success Rates PG1: 0% (0 out of 16 Iterations succeeded). PG2: 40% (6 out of 15 Iterations succeeded). PG3: 100% (4 out of 4 Iterations succeeded). PG4: 40% (4 out of 10 Iterations succeeded). PG5: 100% (5 out of 5 Iterations succeeded). 2. Failure Rates PG1: 100% (16 out of 16 Iterations failed). PG2: 60% (9 out of 15 Iterations failed). PG3: 0% (0 out of 4 Iterations failed). PG4: 60% (6 out of 10 Iterations failed). PG5: 0% (0 out of 5 Iterations failed).

## Discussion

### Participant Prompts and AI-ChatGPT's Effectiveness.

This research investigates how artificial intelligence (AI), particularly AI-ChatGPT, can support computational thinking (CT) skills in mathematical tasks using GeoGebra software, specifically in Measurement and Evaluation courses at the College of Education in Nigeria. The study demonstrates that detailed and structured prompts yield more effective responses from AI-ChatGPT, thereby enhancing students' critical thinking and problem-solving skills. Participants with prior knowledge of GeoGebra performed better in debugging and abstraction tasks than those without experience. Participants, such as PG5, who explicitly defined the task and the AI's role, received more accurate GeoGebra commands compared to PG1, who relied on vague prompts. This finding aligns with Hardman (2023), who emphasises that effective prompt engineering is crucial for maximising the utility of AI tools in education. Moreover, the iterative refinement of prompts observed in PG2 demonstrates the importance of experimentation in improving AI-assisted learning outcomes. Recent studies suggest that incorporating skills into AI literacy training for educators and students is beneficial (Li et al., 2023). Such training could enhance the ability to craft precise instructions, improving AI-ChatGPT's utility across diverse educational contexts.

This contrasts with prior studies such as Yunianto et al., (2024), which reported broader CT coverage in less technical tasks. Limitations of AI-ChatGPT: While ChatGPT facilitated CT development, its inconsistent command accuracy highlights a need for improvement. As Baidoo-Anu and Owusu Ansah (2023) suggest, refining AI training datasets could enhance their performance. This study's findings are critically compared with prior studies, such as Yunianto et al. (2024) on the role of AI-ChatGPT in programming education and Ziatdinov & Valles (2022) on the benefits of GeoGebra's CT. This highlights the study's contributions and limitations. The discussion explicitly addresses limitations, such as ChatGPT’s inconsistent accuracy in GeoGebra commands and its limited adaptability to complex mathematical tasks. Recommendations include teacher training to supplement AI tools and ensure the development of robust CT skills.

**Computational Thinking (CT) Development**

AI-ChatGPT successfully assisted participants with algorithmic thinking and debugging while using GeoGebra, enabling them to create commands and identify errors. However, skills like abstraction and pattern recognition were not as common, especially among those with less experience, like PG1. If tasks become more complex, opinions on AI, such as ChatGPT, might change. Currently, it primarily assists users with basic skills, and further studies are needed to determine if its positive feedback persists across various tasks. Most participants liked using AI-ChatGPT for their math tasks involving programming GeoGebra, suggesting that students could benefit similarly. Users can enhance their understanding by interacting with AI-ChatGPT, especially if they already have a basic understanding. Previous studies support the notion that AI can aid in learning programming, as demonstrated by Ellis & Slade (2023), particularly for GeoGebra objects. Further research is necessary to explore how school students might use AI-ChatGPT for math-related computational thinking tasks. AI technologies are known to improve student performance in various subjects, including mathematics, according to Li et al. (2023) and Plata et al. (2023). This study specifically examines computational thinking in math lessons using GeoGebra, providing evidence of AI-ChatGPT's ability to support skills such as debugging and programming in this context. AI-ChatGPT has created opportunities for learning programming, although its limited ability to generate correct code means that students may still need human support for deeper skill development. Studies by Dos Santos and Cury (2023) emphasize the value of AI-ChatGPT as a virtual peer that gives immediate feedback, helping users with debugging.

Similarly, Yunianto et al. (2024) highlighted how AI can guide students through trial-and-error learning in programming.

**Role of Prior Knowledge**

Participants with experience in GeoGebra or programming (e.g., PG5) performed better in debugging and abstraction tasks, showing that prior knowledge helps the effectiveness of AI-ChatGPT. Some codes generated by AI-ChatGPT were incorrect, often due to unreliable training sources, underscoring the need for improved training data. A lack of understanding of GeoGebra commands and mathematical representations contributed to participants' challenges in creating the required objects. Users with a solid grasp of basic commands were more successful in using AI-ChatGPT effectively. Limitations like restricted understanding and bias in AI-ChatGPT's training data further impacted its performance, as highlighted by Baidoo-Anu and Owusu Ansah (2023). Despite these limitations, some participants successfully trained AI-ChatGPT to generate accurate code. Users should verify AI responses for accuracy, and enhancing their understanding of GeoGebra can improve their interactions with the tool. Data collected from participants revealed that those familiar with GeoGebra had positive experiences with AI-ChatGPT, while beginners found it less helpful. The findings support those of Ellis & Slade (2023), who have shown that while AI tools can enhance learning, their success relies on users' foundational knowledge. This finding corroborates Baidoo-Anu and Owusu Ansah (2023), who argue that while AI tools can enhance learning, their effectiveness depends on users' foundational knowledge and the ability to evaluate AI outputs critically. This suggests that incorporating basic GeoGebra training into curricula could enhance interactions with AI, such as ChatGPT, promoting better learning outcomes.

**Challenges with AI-ChatGPT’s Responses**

While AI-ChatGPT was beneficial, its responses occasionally included errors, such as incorrect syntax or invalid commands, for instance, ("Sector (P, Q,n)"). Participants, such as PG1 and PG4, struggled to debug these errors, highlighting limitations in AI’s current capabilities. The researchers witnessed from the participants’ conversations with GeoGebra that the related code did not work or did not exist. The case of the “Point on Circle” command that did not work has led PG1 to inquire if this command exists. Ultimately, PG1 was unable to fix the Point on Circle command successfully but instead used an alternative command. This finding is consistent with Adel et al., (2024) and Sohail et al. (2023), who noted that generative AI tools are prone to inaccuracies due to their reliance on training data, which may not always align with specific contexts like GeoGebra. Educators must play a critical role in helping students navigate these limitations. By fostering critical thinking and cross-checking AI outputs, teachers can ensure that students do not blindly trust AI-generated solutions (Wollny et al., 2023).

**Implications for Differentiated and Personalised Learning**

The study demonstrates that AI-ChatGPT can facilitate differentiated instruction by adapting to the varying needs of participants. Some learners, such as PG2, received helpful feedback, while others, like PG5, utilised AI as a peer tutor to enhance their skills. Differentiated instruction involves five key areas: content, instructional strategies, the classroom, products, and instructors (Reis & Renzulli, 2018). Teachers can adjust any of these areas to cater to individual learning needs and enhance their understanding of Mathematics. Using AI-ChatGPT allows for varied prompts, even if users are working on the same tasks, which helps in differentiated instruction as noted by Li et al. (2023). This variation arises from the use of different instructional strategies. AI-ChatGPT offer personalised feedback and creates interactive learning experiences, allowing for a tailored learning path. Dos Santos and Cury (2023) support this view, emphasising that AI can provide learning experiences that cater to individual strengths and challenges. However, to achieve true personalization, AI tools should be integrated into a broader pedagogical framework that includes explicit instruction and teacher support.

## Limitations and Future Directions

Technical limitations: The inconsistent accuracy of AI-ChatGPT highlights the need to refine training datasets to include domain-specific contexts such as GeoGebra commands as suggested by (Baidoo-Anu & Owusu Ansah, 2023). Broader aspects of CT: Future research should investigate how AI tools can better support underexplored aspects of CT, such as abstraction and generalisation, in more complex tasks, which aligns with the opinion of Shute et al. (2017). Educational policy integration: Policymakers should explore how to integrate AI knowledge into teacher education programs and curricula to prepare students for AI-enabled learning environments, as noted by Li et al. (2023).

## Conclusion and Recommendations

This study highlights the potential of integrating AI (ChatGPT) with GeoGebra to enhance computational thinking in mathematics education. The results demonstrate the usefulness of ChatGPT for facilitating algorithmic thinking and error correction, although limitations in abstraction and consistent precision of commands remain. Broader measures, such as requiring policymakers to consider integrating AI tools like ChatGPT into mathematics curricula to promote computational thinking, are required. Teacher training programs should incorporate AI-assisted instruction to equip teachers with the skills necessary to guide students effectively. Highlight future research avenues, such as examining AI performance in other mathematical domains and modifying AI training datasets to improve the accuracy of GeoGebra training. It is recommended that:

1. Create standard guidelines and test beds to optimize AI-ChatGPT responses.
2. Modify the AI training dataset to improve the accuracy of generating mathematical commands.

## References

Adel, A. Ashan,A. & Davidson, C., (2024). ChatGPT promises and challenges in Education: Computational and Ethical perspectives. Educational Science, 14(8), 814.

Ali, A. (2024). ChatGPT Needs SPADE (Sustainability, PrivAcy, Digital divide and Ethics A Review. xrXiv preprint arXiv: 2305.03123.

Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI):Understanding the potential benefits of AI-ChatGPT in promoting teaching and learning. *Journal of AI, 7*(1), 52-62. https://doi.org/10.61969/jai.1337500

Botana, F., Recio, T., & Velez, M. P. (2024). " On Using GeoGebra and ChatGPT for Geometric Discovary." Computers, 13(8),187

Cascella, M., Montomoli, J., Bellini, V., & Bignami, E. (2023). Evaluating the feasibility of AI-ChatGPT in healthcare: An analysis of multiple clinical and research scenarios. *Journal of Medical Systems,47*(1), 33.https://doi.org/10.1007/s10916-023-01925-4

Castelvecchi, D. (2023). How will AI change mathematics? Rise of chatbots highlights discussion. *Nature, 615*(7950), 15-16. https://doi.org/10.1038/d41586-023- 00487-2

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of AI-ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Dos Santos, O. L., & Cury, D. (2023). Challenging the confirmation bias: Using AI-ChatGPT as a virtual peer for peer instruction in computer programming education. *2023 IEEE Frontiers in Education Conference (FIE)* (pp. 1–7).

Ellis, A. R., & Slade, E. (2023). A new era of learning: Considerations for ChatGPT as a tool to enhance statistics and data science education. *Journal of Statistics and Data Science Education, 31*(2), 128–133. https://doi.org/10.1080/26939169.2023.2223609

Hardman, J. (2023). Crafting effective prompts for AI learning tools: A framework for educators. *Journal of Educational Technology*, 10(4), 123-140.

Li, X., Zhang, Y., & Chen, H.(2023).Integrating AI tools in STEM education: Opportunities and challenges. International Journal of STEM Education, 5(3), 202-220.

Olivares Funes, J., Droguett Parads, B., Martin de Julian, P. And Burgos Villegas, A. (2024). Incorporation of GeoGebra Applets and ChatGPT to Strenthengthen the learning of Exact Equations in Engineering Students.' *Mathematics and Computer Science Development*, 7.133-143.

Papert, S. (1980). *Mindstorms: Children, Computers, and Powerful Ideas* (1st ed.). Basic Books.

Plata, S., De Guzman, M. A., & Quesada, A. (2023). Emerging research and policy themes on academic integrity in the age of Chat GPT and generative AI. *Asian Journal of University Education, 19*(4), 743–758. https://doi.org/10.24191/ajue.v19i4.24697

Rane, N.(2023).Enhancing mathematical capabilities through AI-ChatGPT and similar generative artificial intelligence: Roles and challenges in solving mathematical problems. *SSRN Electronic Journal.*

Reis, S. M., & Renzulli, J. S. (2018). The five dimensions of differentiation. *International Journal for Talent Development and Creativity, 6*(1), 87-94. https://files.eric.ed.gov/fulltext/EJ1296874.pdf

Rudolph, J., Tan, S., & Tan, S.(2023).War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. Journal of Applied Learning and Teaching. Sohail, S.S., et al.(2023). Decoding ChatGPT: A taxonomy of existing research, current challenges, and possible future directions. Journal of King Saud University - Computer and Information Sciences.

Shabunina, V., Sarancha, V., Maslak, V., Shevchenko, O., & Tur, O. (2023). Educational potential of ChatGPT: Teaching tool for students' competencies development. *2023 IEEE 5th International Conference on Modern Electrical and Energy System (MEES)* (pp. 1–6). https://doi.org/10.1109/MEES61502.2023.10402380

Shute, V.J., Sun, C., & Asbell-Clarke, J.(2017). Demystifying computational thinking. Educational Research Review, 22, 142-158.https://doi.org/10.1016/j.edurev.2017.09.003

Sohail, S., Yunianto, W., & Prodromou, T.(2023). Decoding AI-ChatGPT: A taxonomy ofexisting research, current challenges, and possible future directions. Journal of Educational AI, 12(2), 45-62.

van Borkulo, S. P., Chytas, C., Drijvers, P., Barendsen, E., & Tolboom, J. (2021). Computational thinking in the mathematics classroom: Fostering algorithmic thinking and generalization skills using dynamic mathematics software. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3481312.3481319.

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

Wollny, S., Rosenberg, D., & Scherer, R.(2023).Trust in generative AI: Exploring students' reliance on ChatGPT for STEM education. Computers & Education, 190, 104620

Yunianto, W., Prodromou, T., & Lavicza, Z. (2023). Debugging on GeoGebra-based Mathematics +Computational Thinking lessons. *28th Asian Technology Conference in Mathematics* (pp. 322-331). https://atcm.mathandtech.org/EP2023/regular/22065.pdf

Yunianto, W., Galic, S., & Lavicza, Z. (2024).Exploring Computational Thinking in Mathematics Education: Integrating ChatGPT with GeoGebra for Enhanced learning Experiences. International Journal of Education in Mathematics, Science and Technology, 12(6), 1451-1470.

Ziatdinov, R & Valles, J. R.(2022) Synthesis of modeling, visualization and programming in GeoGebra as an effective approach for teaching and learning STEM Topics Mathematics 10(3),398.