International Journal of Multicultural and Multireligious Understanding

http://ijmmu.com
editor@ijmmu.com
ISSN 2364-5369
Volume 12, Issue 12
December, 2025
Pages: 40-48

Fostering Mathematical Conceptual Understanding through ChatGPT-Assisted Discovery Learning

Dian Nur Utami; Sugiman

Department of Mathematics Education, Yogyakarta State University, Indonesia

http://dx.doi.org/10.18415/ijmmu.v12i12.7200

Abstract

This study examines the impact of the Discovery Learning model assisted by ChatGPT on students' mathematical conceptual understanding. Employing a quasi-experimental Non-Equivalent (Pretest–Posttest) Control Group design, two classes participated: one received instruction through Discovery Learning integrated with ChatGPT, while the control group was taught using Direct Instruction. The conceptual understanding test, validated and confirmed reliable, assessed five indicators of mathematical understanding. The results revealed that the experimental group achieved significantly higher posttest scores and greater normalized gains (N-Gain = 0.212) compared to the control group (N-Gain = 0.131). Notably, the most substantial improvement occurred in transforming mathematical representations. These findings suggest that incorporating ChatGPT within Discovery Learning fosters deeper conceptual understanding through adaptive feedback, guided reasoning, and reflective engagement, highlighting its potential as an effective pedagogical innovation for mathematics education in the digital era.

Keywords: *ChatGPT; Discovery Learning; Conceptual Understanding; Mathematics Education*

Introduction

The emergence of Generative Artificial Intelligence (GenAI) has introduced fresh prospects within the realm of education. UNESCO (2023) officially recommends the use of GenAI in educational settings through its document *Guidance for Generative AI in Education and Research*. This technology provides new possibilities for adaptive, interactive, and personalized learning. One notable example of GenAI is the Chat Generative Pretrained Transformer (ChatGPT), developed by OpenAI. ChatGPT functions as a language model engineered to process directives through an interactive conversational interface, facilitating ongoing dialogue, identifying inaccuracies, declining unsuitable inquiries, and rectifying erroneous premises (OpenAI, 2022).

Several studies have indicated that ChatGPT holds significant potential to support educational processes, particularly as an adaptive and interactive learning tool (Dahlan et al., 2024; Kasneci et al., 2023; Kiryakova & Angelova, 2023). In the context of mathematics education, its implementation has been shown to facilitate students' understanding of complex mathematical concepts (Ramprakash et al.,

2024) and to enhance their problem-solving skills (Yunianto et al., 2024). Nevertheless, some researchers have identified limitations in the use of ChatGPT. Dasari et al. (2024) emphasize that ChatGPT is not yet fully effective in fostering deep conceptual understanding, critical thinking, and guidance that resembles human interaction. Furthermore, Wardat et al. (2023) highlight that the accuracy of ChatGPT's responses largely depends on the quality of the prompts or instructions provided by users.

Mathematical conceptual understanding is a fundamental aspect of mathematics learning, as it serves as the foundation for mastering higher-order mathematical skills. Earlier research has indicated a positive association between conceptual comprehension and various other mathematical proficiencies. Kharis et al. (2021) found that conceptual understanding has a significant effect on mathematical literacy, while Kholid et al. (2021) demonstrated that students with a high level of conceptual understanding are more likely to succeed in mathematical problem-solving. However, several studies have shown that high school students continue to face difficulties in understanding mathematical concepts and generally exhibit low levels of conceptual comprehension (Cahyani et al., 2024; Lumbantoruan & Manalu, 2024; Muin & Fatma, 2021; Sasmita & Qohar, 2021). This condition underscores the importance of instructional strategies that can enhance students' conceptual understanding, particularly through the use of innovative educational technologies such as ChatGPT.

One approach that has the potential to address the limitations of ChatGPT is the Discovery Learning model, which emphasizes students' active involvement in discovering concepts, rules, or relationships through exploration. Glaser (1966) stated that Discovery Learning is an instructional process that involves teaching relationships, concepts, or principles. Learning through the Discovery Learning model encourages active participation and student engagement, which can help learners develop deeper cognitive skills (Stoffova, 2020). The steps of the Discovery Learning model applied in this study are as follows: (1) Introduction, to stimulate curiosity and prepare students to face the problem; (2) Developing Understanding, where students explore and identify patterns or conceptual relationships; (3) Verification, a crucial stage to test the validity and consistency of their findings; and (4) Application, applying the discovered concepts in new contexts to deepen understanding.

The implementation of ChatGPT integrated within the Discovery Learning model is expected to minimize potential biases in students' mathematical conceptual understanding. In the introduction stage, ChatGPT can stimulate students' curiosity through adaptive and inquiry-driven prompts. During the developing understanding phase, ChatGPT provides access to extensive information and supports students in exploring mathematical concepts. The verification stage is crucial, as it addresses ChatGPT's limitations in providing information that may not always be accurate. Through teacher-guided verification, students are encouraged to examine, confirm, and evaluate the validity of the information generated by the AI. Consequently, students are not merely passive recipients but active critics of their own exploratory learning process. In the application stage, ChatGPT can serve as a learning assistant that helps students apply the acquired concepts in various contextual situations. This integrative approach aligns with the findings of Sudirman & Rahmatillah (2023), who revealed that AI-based tools such as ChatGPT can enrich Discovery Learning by making the learning process more interactive and engaging. Similarly, Huynh et al. (2024) highlighted that AI can facilitate Guided Discovery Learning, in which students explore topics with scaffolded guidance from teachers—an approach proven effective in improving learning outcomes.

Despite its potential, research examining the effects of integrating ChatGPT within the Discovery Learning model on cognitive aspects, such as students' mathematical conceptual understanding, remains limited. Most previous studies have focused either on the general potential of ChatGPT or the effectiveness of Discovery Learning without combining the two, leaving a gap in understanding how AI can specifically enhance students' cognitive abilities. This gap underscores the need for empirical studies to explore how ChatGPT can optimally support the Discovery Learning process in the context of mathematics education.

Based on this background, the present study aims to investigate the effects of Discovery Learning assisted by ChatGPT on students' mathematical conceptual understanding at the senior high school level. The study is designed to evaluate the extent to which the integration of ChatGPT as a digital learning tool can enhance deep conceptual understanding. Practically, this research is expected to guide teachers in designing innovative AI-based instructional strategies that can improve students' conceptual comprehension. Theoretically, it is anticipated to enrich the literature on AI integration in discovery-based learning models, offer insights into optimizing human–machine interactions in educational settings, and serve as a foundation for the development of adaptive and effective learning models in the digital era.

Method

This study employed a quasi-experimental research design using the pretest–posttest control group design. The research aimed to examine the effect of integrating ChatGPT into the Discovery Learning model on students' mathematical conceptual understanding. Two groups participated: an experimental group taught using ChatGPT-enhanced Discovery Learning, and a control group taught through Direct Instruction. Both groups completed a pretest before and a posttest after the intervention to measure learning gains.

Participants

The participants were eleventh grade students from during the 2025/2026 academic year. Two classes were selected through simple random sampling to ensure the comparability of student characteristics. The experimental group consisted of 28 students, while the control group comprised 22 students. Both classes were taught by the same mathematics teacher to minimize instructional bias and maintain consistent teaching conditions throughout the study.

Instruments

The instrument used in this study was a conceptual understanding test in mathematics, developed based on the targeted learning objectives. The test consisted of multiple-choice items designed to measure students' ability to identify, connect, and apply mathematical concepts accurately.

Based on previous studies on indicators of conceptual understanding, the indicators of mathematical conceptual understanding used in this research (Anderson & Krathwohl, 2001; BNSP, 2006; Kilpatcrick et al., 2001; NCTM, 2000; Romberg & Kaput, 1999):

1. Restating a mathematical concept,
2. Identifying examples and non-examples of a mathematical concept,
3. Transforming one form of mathematical representation into another,
4. Applying mathematical concepts, and
5. Connecting various mathematical concepts.

Before being used, the instrument underwent expert validation to ensure content validity. The expert validation results indicated that the instrument was valid for measuring students' mathematical conceptual understanding. A pilot test was also conducted on a similar group of students to establish reliability, and the reliability coefficient calculated using Cronbach's Alpha was 0.678, which is considered acceptable for educational research. Items that did not meet the validity and reliability criteria were revised or removed prior to data collection.

Data Analysis

The data collected from the pretest and posttest were analyzed using statistical inferential methods. Prior to hypothesis testing, the data were examined for normality using the Shapiro Wilk test and homogeneity of variances using Levene's test.

To determine the difference in posttest scores between the experimental and control groups, an independent sample t-test was performed at a significance level of $\alpha = 0.05$.

The interpretation criteria were as follows:

If $p < 0.05$, there is a significant difference between the two groups, indicating that ChatGPT-assisted Discovery Learning has a significant impact on students' conceptual understanding.

If $p \geq 0.05$, there is no statistically significant difference between the two groups.

Results

Descriptive Analysis

The descriptive statistics of students' mathematical conceptual understanding in both the experimental and control classes are presented in Table 1.

Table 1. Descriptive Statistics of Conceptual Understanding Test

Description	Experimental		Control	
	Pretest	*Posttest*	*Pretest*	*Posttest*
Mean	53,32	64,82	38,50	48,86
Standard Deviation	13,34	13,05	16,80	17,12
Theoretical Maximum Score	100	100	100	100
Theoretical Minimum Score	0	0	0	0
Maximum Score	73,00	87,00	73,00	80,00
Minimum Score	27,00	33,00	13,00	20,00

The descriptive results indicate an overall improvement in both groups; however, the experimental class demonstrated a greater increase in mean scores ($\Delta = 11.5$) compared to the control class ($\Delta = 10.36$). This suggests that the Discovery Learning model assisted by ChatGPT had a more substantial impact on students' conceptual understanding.

Item-Based Indicator Analysis

Further analysis by indicator (Table 2) revealed consistent improvement across all conceptual understanding dimensions.

Table 2. Average Scores by Conceptual Understanding Indicator

Indicator	Experimental		Control	
	Pre	Post	Pre	Post
Restating mathematical concepts	0.821	0.910	0.636	0.704
Identifying examples and non-examples	0.500	0.536	0.250	0.568
Changing representation forms	0.452	0.845	0.409	0.561
Connecting mathematical concepts	0.428	0.500	0.386	0.341
Applying mathematical concepts	0.571	0.571	0.307	0.432

The highest increase was observed in the indicator changing representation forms, followed by restating mathematical concepts. The improvement across indicators was more pronounced in the experimental group than in the control group.

Normalized Gain (N-Gain) Analysis

The normalized gain (N-Gain) results indicated that the experimental group achieved a higher average gain score (**0.212**) compared to the control group (**0.131**). Both are categorized as low improvement, yet the experimental group demonstrated a relatively stronger effect of the learning intervention.

Normality and Homogeneity Tests

Shapiro Wilk tests confirmed that the data for both groups were normally distributed ($p > 0.05$). Levene's test also showed that variances were homogeneous ($p > 0.05$), indicating that the data met the assumptions required for parametric analysis.

Hypothesis Testing

The independent sample t-test results are presented in Table 3.

Table 3. Independent Sample T-Test Results

Test	Statistic	p-value	Decision
Independent Sample T-Test	3.499	< 0.001	H_0 rejected

The results show a significant difference between the experimental and control groups ($p < 0.001$). This indicates that students who received instruction using the Discovery Learning model assisted by ChatGPT achieved significantly higher conceptual understanding scores than those taught using the Direct Instruction model.

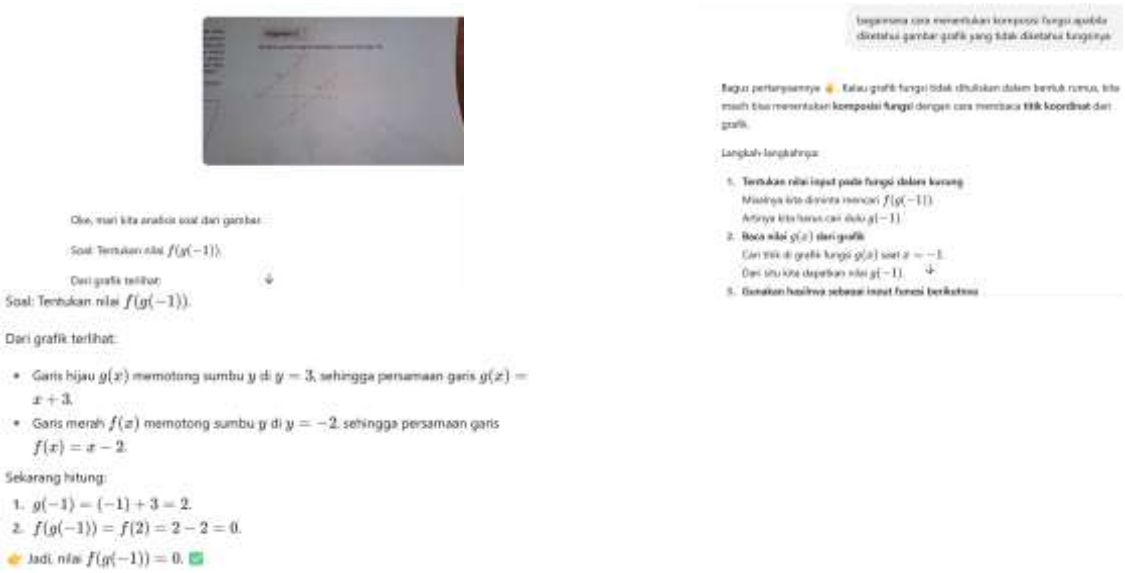
Discussion

The findings of this study indicate that the application of the Discovery Learning model assisted by ChatGPT had a positive effect on students' mathematical conceptual understanding. Students in the experimental group achieved significantly higher posttest scores than those in the control group, suggesting that the integration of AI tools such as ChatGPT can enhance the effectiveness of inquiry-based learning environments.

Further analysis based on conceptual understanding indicators revealed consistent improvement across all aspects. The highest mean scores in both groups were found in the first indicator (restating a mathematical concept), whereas the lowest were in connecting multiple mathematical concepts (experimental class) and identifying examples and non-examples (control class).

A notable improvement occurred in the third indicator, namely transforming one mathematical representation into another from 0.452 to 0.845 in the experimental group, compared to 0.409 to 0.561 in the control group. This suggests that ChatGPT effectively supported students in visualizing and connecting mathematical representations.

ChatGPT.



Images 1. Study Case

An example can be seen during a learning activity on function composition (Image 1), where students interacted with ChatGPT to solve problems from the worksheet. Initially, students uploaded diagrams without providing sufficient context, leading to incorrect AI responses. After teacher guidance, they refined their prompts, enabling ChatGPT to generate accurate explanations and reasoning steps. This process encouraged students to think critically about the relationships among mathematical elements and improved their ability to represent and interpret concepts meaningfully. Thus, prompt refinement became a reflective thinking process that strengthened students' conceptual understanding.

For the fourth indicator (connecting mathematical concepts), the experimental group improved from 0.428 to 0.500, whereas the control group slightly decreased from 0.386 to 0.341. This finding implies that Gen-AI-supported Discovery Learning helps students establish conceptual links across topics through adaptive examples and feedback.

Although the fifth indicator (applying concepts in problem solving) showed a relatively stable score in the experimental group (0.571), their performance remained higher than that of the control group. This suggests that while students had solid conceptual foundations, they still needed further guidance to transfer their understanding into higher-order problem-solving contexts.

This improvement aligns with previous studies (Dahlan et al., 2024; Kasneci et al., 2023) which emphasize that ChatGPT promote interactive and adaptive learning experiences. Within the Discovery Learning framework, ChatGPT functioned as an intelligent scaffolding tool, supporting students during exploration and conceptualization phases. The introduction, developing understanding, verification, and application stages encouraged students to actively engage with mathematical problems, construct meaning, and verify their conceptual understanding through AI-assisted dialogue.

A particularly important aspect observed in this study was the verification stage, where students cross-checked their reasoning and answers using ChatGPT. This process helped mitigate one of ChatGPT's known limitations, its tendency to produce plausible but inaccurate responses (Dasari et al., 2024). By combining teacher guidance with AI-based feedback during verification, students were encouraged to critically evaluate responses, enhancing conceptual accuracy and critical thinking.

The results also support the theoretical foundations of Discovery Learning proposed by Glaser (1966) and Bruner (1961), both of whom emphasize learner autonomy and active cognitive engagement in the construction of knowledge. The significant improvement in the indicators of "restating mathematical concepts" and "changing representation forms" suggests that students not only recalled concepts but were also able to reinterpret and connect them across various contexts. This finding aligns with the core idea of Discovery Learning, that meaningful understanding develops when learners are actively involved in constructing their own knowledge rather than passively receiving information. In this process, ChatGPT functions as a digital scaffold, providing dynamic examples, adaptive feedback, and multiple representations that help make abstract mathematical concepts more concrete and accessible. These results are consistent with the findings of Kholid et al. (2021) and Kharis et al. (2021) who reported that conceptual understanding is closely related to mathematical literacy and problem-solving ability.

Although the average normalized gain (N-Gain = 0.212) was categorized as low, the result still demonstrates meaningful progress compared to the control class (N-Gain = 0.131). This modest gain may be attributed to the relatively short intervention duration and the novelty of AI-assisted learning for both teachers and students. Nevertheless, the improvement pattern indicates that consistent and well-structured integration of ChatGPT could produce stronger learning outcomes over time.

Overall, this study provides empirical evidence that integrating ChatGPT into Discovery Learning can enhance students' conceptual understanding of mathematics by promoting active exploration, guided reasoning, and reflective verification. These findings highlight the pedagogical potential of ChatGPT as a supportive tool rather than a replacement for human instruction, emphasizing the importance of teacher mediation in maintaining conceptual depth and accuracy in AI-assisted learning environments.

Conclusion

This study found that the Discovery Learning model assisted by ChatGPT significantly improved students' mathematical conceptual understanding. Students in the experimental group achieved higher posttest scores, especially in restating concepts and transforming representations, indicating deeper conceptual processing supported by AI-assisted exploration.

ChatGPT acted as adaptive scaffolding during the exploration and verification stages, helping students critically evaluate and refine their reasoning. This aligns with Bruner's and Glaser's theories of Discovery Learning, emphasizing active knowledge construction and learner autonomy.

Although the normalized gain (N-Gain = 0.212) was low, it showed meaningful progress compared to the control group. Overall, integrating ChatGPT within Discovery Learning effectively enhances conceptual understanding by promoting active exploration, guided reasoning, and reflective verification in mathematics learning.

Recommendations

Based on the findings, several recommendations can be made. First, teachers are encouraged to integrate Generative AI tools, such as ChatGPT, within Discovery Learning to facilitate active exploration and verification of mathematical concepts. Teacher guidance remains essential to ensure that AI-generated responses are critically evaluated and conceptually accurate. Second, future research should extend the duration and scope of AI-assisted interventions to examine long-term effects on higher-order thinking skills and problem-solving abilities. Comparative studies across different mathematical topics and educational levels are also recommended to strengthen the generalizability of these findings. Finally, developers and policymakers should collaborate to design AI-integrated learning environments that support personalized feedback, ethical use, and pedagogical alignment with inquiry-based learning models.

References

Anderson, L. W., & Krathwohl, D. R. (2001). A Taxonomy for Learning, Teaching and Assessing, a Bridged Edition. *Pearson Education*, *51*(275), 560.

BNSP. (2006). *Model Penilaian Kelas*. Depdiknas.

Bruner, J. S. (1961). The Act of Discovery. *Havard Educational Review*, *31*, 21–32.

Cahyani, W. M., Fathia, Y., Wulandari, W., Kulsum, Z. N., Hafiz, M., & Putri, F. M. (2024). Analysis of Student Concept Understanding Ability on Absolute Value Rational and Irrational Inequality Materials. *AIP Conference Proceedings*, 060032. https://doi.org/10.1063/5.0202015.

Dahlan, M. A., Omar, R., & Kamarudin, S. (2024). ChatGPT in Academia: A Holistic Examination of Student Performance and Future Implications. In *Business Development via AI and Digitalization* (Vol. 538, pp. 617–627). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-031-62102-4_51.

Dasari, D., Hendriyanto, A., Sahara, S., Suryadi, D., Muhaimin, L. H., Chao, T., & Fitriana, L. (2024). ChatGPT in Didactical Tetrahedron, Does it Make an Exception? A Case Study in Mathematics Teaching and Learning. *Frontiers in Education*, *8*. https://doi.org/10.3389/feduc.2023.1295413.

Glaser, R. (1966). Variables in Discovery Learning. In L. S. Shulman & E. R. Keislar (Eds.), *Learning by Discovery: A Critical Appraisal*. Rand Mcnally & Company.

Huynh, C. N. T., Van Thi Tran, A., Bui, T. T., Nguyen, H. T., & Tran, P. V. (2024). Applying Guided Discovery Learning to Enhance the Achievement of Information Technology Team. In *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering* (Vol. 579, pp. 186–196). Springer, Cham. https://doi.org/10.1007/978-3-031-58878-5_14.

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. *Learning and Individual Differences*, *103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274.

Kharis, S. A. A., Salsabila, E., & Haeruman, L. D. (2021). Effect of Mathematical Concept Understanding and Mathematical Reasoning on Mathematical Literacy Abilities. *Journal of Physics: Conference Series*, *1747*(1), 012042. https://doi.org/10.1088/1742-6596/1747/1/012042.

Kholid, M. N., Imawati, A., Swastika, A., Maharani, S., & Pradana, L. N. (2021). How are Students' Conceptual Understanding for Solving Mathematical Problem? *Journal of Physics: Conference Series*, *1776*(1), 012018. https://doi.org/10.1088/1742-6596/1776/1/012018.

Kilpatcrick, J., Swafford, J., & Findell, B. (2001). *Adding + It Up: Helping Children Learn Mathematics*. National Academy Press.

Kiryakova, G., & Angelova, N. (2023). ChatGPT—A Challenging Tool for the University Professors in Their Teaching Practice. *Education Sciences*, *13*(10), 1056. https://doi.org/10.3390/educsci13101056.

Lumbantoruan, J. H., & Manalu, R. U. (2024). Effectiveness of Learning Mathematics Derivative Materials Using Modules Equipped with Cooperative Models in High Schools. *International Journal of Evaluation and Research in Education (IJERE)*, *13*(1), 523. https://doi.org/10.11591/ijere.v13i1.25354.

Muin, A., & Fatma, M. (2021). Hypothetical Learning Trajectory Design in Development of Mathematics Learning Didactic Design in Madrasah. *Journal of Physics: Conference Series*, *1836*(1), 012070. https://doi.org/10.1088/1742-6596/1836/1/012070.

NCTM. (2000). *Principles and Standards for School Mathematics*. Inc.

OpenAI. (2022, November 30). *Introducing ChatGPT*. Https://Openai.Com/Index/Chatgpt/. https://openai.com/index/chatgpt/.

Ramprakash, B., Surya, D. B., Nithyakala, G., Bhumika, K., & Avanthika, S. (2024). Comparing Traditional Instructional Methods to ChatGPT: A Comprehensive Analysis. *Journal of Engineering Education Transformations*, *37*.

Romberg, T. A., & Kaput, J. J. (1999). Setting The Stage. In E. Fennema & T. A. Romberg (Eds.), *Mathematics Classroom That Promote Understanding*. Lawrence Erlbaum Associates, Inc.

Sasmita, S. A., & Qohar, A. (2021). Implementation of problem based learning to improve students' understanding of systems of linear equations in three variables. *Journal of Physics: Conference Series*, *1918*(4), 042055. https://doi.org/10.1088/1742-6596/1918/4/042055.

Stoffova, V. (2020). Discovery Learning by Interactive Animation Models. *ELearning and Software for Education Conference*, 246–252. https://doi.org/10.12753/2066-026X-20-116.

Sudirman, I. D., & Rahmatillah, I. (2023). Artificial Intelligence-Assisted Discovery Learning: An Educational Experience for Entrepreneurship Students Using ChatGPT. *2023 IEEE World AI IoT Congress (AIIoT)*, 0786–0791. https://doi.org/10.1109/AIIoT58121.2023.10174472.

UNESCO. (2023). *Guidance for Generative AI in Education and Research*. UNESCO. https://doi.org/10.54675/EWZM9535.

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.

Yunianto, W., Lavicza, Z., Kastner-Hauler, O., & Houghton, T. (2024). Investigating the Use of ChatGPT to Solve a GeoGebra Based Mathematics+Computational Thinking Task in a Geometry Topic. *Journal on Mathematics Education*,*15*(3),1027–1052.https://doi.org/10.22342/jme.v15i3.pp1027-1052.