Jurnal Elemen, 11(2), 297-310, April 2025
https:/doi.org/10.29408/jel.v11i2.27842

e- ISSN: 2442 - 4226

Measuring the impact of ChatGPT in enhancing number sense acquisition: A mixed-method study

Kadir [1 *], Gita Fajrin Jafar [1], Benjamin B. Mangila [2]

[1] Faculty of Education and Teacher Training, Manado State Islamic Institute, North Sulawesi, Indonesia
[2] Josefina H. Cerilles State College, Mindanao, Philippines

[*] Correspondence: kadir@iain-manado.ac.id

Abstract

The effectiveness of ChatGPT in enhancing mathematics learning outcomes for students remains a topic of debate despite its growing application in educational settings. This study seeks to assess the impact of ChatGPT on students' understanding of number concepts in mathematics, particularly for 7th-grade junior high school students, by employing a combination of classroom experiments and a control group while gathering students' perspectives on its use. In a mixed-method approach, the quantitative findings indicate that ChatGPT usage does not significantly improve students' academic performance in number concepts. However, qualitative insights from the classroom experiments reveal that students perceived several benefits, such as increased efficiency in studying, better comprehension of number-related problems, enhanced accuracy in solving problems, and assistance in overcoming learning challenges. While ChatGPT did not directly affect academic outcomes, it was found to be a valuable supplementary tool in the learning process. Based on these findings, this study recommends that educators consider incorporating ChatGPT to complement traditional teaching methods to enhance the overall learning experience of number concepts.

Keywords: ChatGPT; learning outcomes; mathematics; students

How to cite: Kadir, Jafar, G. F, & Mangila, B. B. (2025). Measuring the impact of ChatGPT in enhancing number sense acquisition: A mixed-method study. *Jurnal Elemen*, *11*(2), 297-310. https://doi.org/10.29408/jel.v11i2.27842

Received: 18 October 2024 | Revised: 10 Januari 2025
Accepted: 11 April 2025 | Published: 29 May 2025

Introduction

Numbers are one of the fundamental topics in mathematics, playing a crucial role as the foundation for various mathematical concepts, from basic operations to more complex calculations (Charles & Carmel, 2005; Resnick, 2020). However, in practice, students often struggle to develop a deep understanding of numbers (Sowder, 2020; Zalima et al., 2020). These difficulties are often attributed to conventional teaching methods emphasizing memorization and mechanical procedures rather than fostering strong conceptual understanding (Wijaya et al., 2015). Studies have shown that traditional approaches, such as lectures and routine exercises (Hemmi et al., 2019), tend to be rigid, less interactive, and less effective in enhancing students' critical thinking and problem-solving skills (Abah, 2020).

One way to overcome difficulties in understanding number concepts, which are believed to stem from traditional learning methods, is to utilize modern and relevant technology. Technology has become an integral component of educational processes. Artificial intelligence, such as ChatGPT, is an increasingly utilized innovation. The selection of this AI model over others is based on its various promising positive potentials in learning, as indicated by previous literature. ChatGPT can aid in simplifying complex concepts, answering questions, and offering comprehensive explanations throughout the learning process (Leelavathi & Surendhranatha, 2024; Tang et al., 2024). In mathematics, a subject often perceived as challenging, introducing this technology is expected to enhance learning outcomes by providing additional support in understanding key concepts (Alneyadi & Wardat, 2023; Goos et al., 2020; Rudolph et al., 2023).

The various advantages of ChatGPT that have been explained previously are worthy of being used as a potential tool to support learning number concepts, which has experienced obstacles so far. With its ability to generate diverse explanations, present contextual examples, and provide instant feedback (Rane, 2023), ChatGPT may help students better understand number concepts. Several studies reveal that AI, such as ChatGPT, can enhance students' mathematical skills, problem-solving abilities (Alneyadi & Wardat, 2023; Wardat et al., 2023), and critical thinking (Sánchez-Ruiz et al., 2023).

However, despite the many benefits offered, the effectiveness of ChatGPT in improving learning outcomes remains a subject of debate. Remoto (2024) found that AI can sometimes produce inaccurate answers, which poses a risk of creating misconceptions in understanding numbers. Additionally, Xue et al. (2024) found no significant impact of ChatGPT usage on students' academic skills. Given the limited and inconsistent research findings, further studies are needed to explore the benefits and challenges of using ChatGPT in number concept learning.

Several studies claim that ChatGPT can assist students in independently exploring number concepts (Frieder et al., 2023) and solving number problems more efficiently (Cotton et al., 2023; Rane, 2023). However, a critical question arises regarding the extent to which this technology can truly enhance student learning outcomes in a significant way. Despite its promising potential, there is still insufficient evidence to suggest that ChatGPT can replace traditional learning methods or directly improve students' understanding of number concepts.

The scarcity of research on the effectiveness of ChatGPT in mathematics education, particularly in number concepts, creates a gap that needs to be addressed. Information on the

impact of ChatGPT in number concept learning can contribute to the development of literature by clarifying the role of ChatGPT in supporting mathematics learning outcomes and a comprehensive analysis of the strengths and weaknesses of ChatGPT in number concept learning. The findings of this research can serve as a reference for educators in maximizing the benefits of ChatGPT while mitigating its potential risks.

Methods

This research employed a mixed-method approach, integrating both quantitative and qualitative data within a single study (Bergman, 2008), to evaluate student learning outcomes in mathematics and analyse students' perceptions of using ChatGPT. Specifically, a parallel mixed-method research design was implemented, wherein quantitative and qualitative data were collected and analysed concurrently, with interpretations derived from the combined results of both approaches (Creswell et al., 2008).

This research also employed a learning intervention by adopting a quasi-experimental design, as randomly assigning students to the experimental and control groups would have been impractical (Campbell & Stanley, 2015). The study was conducted at Junior Highscholl State 10 Manado due to the school's consent to participate, ease of data collection, and the availability of sufficient technological infrastructure, including internet access and hardware required to use ChatGPT. Additionally, the selection of two classrooms (experimental group and control group) was based on the school's resources. Classroom selection considered equal student ability to minimize research bias. The experimental group received instruction incorporating ChatGPT, while the control group received conventional teaching.

The research began with students completing a pre-test. Both the experimental and control groups participated in seven study sessions, each lasting 90 minutes. In the experimental group, instruction was designed with the assistance of ChatGPT. This group experienced four distinct types of sessions: (1) an instruction session, where the teacher used ChatGPT to provide additional explanations for concepts students found difficult to understand; (2) a question-and-answer session, in which students posed mathematics-related questions to ChatGPT and received accurate, relevant answers; (3) a problem-solving session, where the teacher assigned math problems, allowing students to pair up and use ChatGPT as a tool for problem-solving; and (4) an independent exercise session, where students were given additional tasks to complete with ChatGPT as a support resource. In contrast, the control group followed conventional teaching methods as regularly practiced by the school. Upon completion of the study period, students in both groups took a post-test and submitted a short essay.

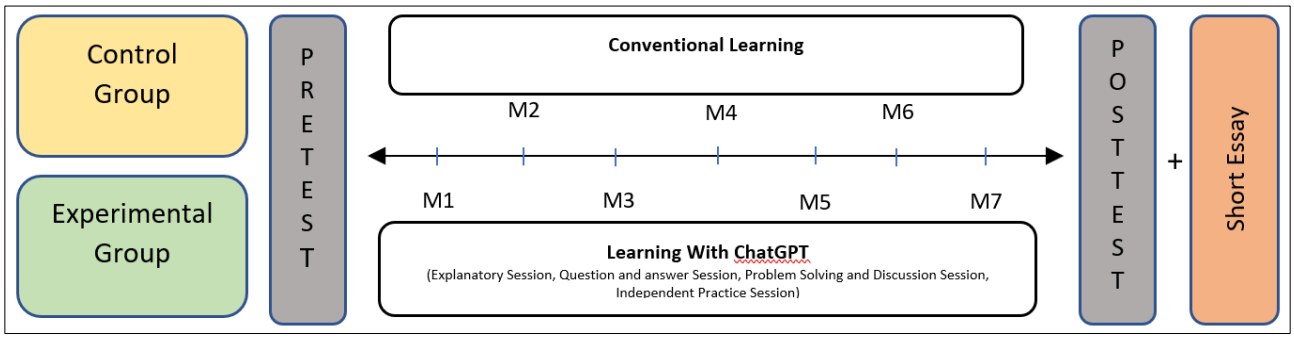

Figure 1. Research procedure

The research subject was students in the 7th grade from Junior High School State 10 Manado, the first semester of the 2024–2025 academic year. Two classrooms were selected; one was the experiment group, which received instruction utilizing ChatGPT (19 male and 20 female students), and the other was the control group with traditional teaching methods (19 male and 18 female students).

In collecting quantitative data, the research used tests in the form of a math test. This test contained 20 multiple-choice questions categorized as HOTS (n=11) and LOTS (n=9). Correct answers were awarded 5 points, while incorrect answers received 0 points, resulting in a score range from 0 to 100. To confirm the validity of the test, experts in evaluation, measurement, and mathematics were involved in providing reviews and feedback, leading to revisions and modifications to produce the final version. Given this process, the instrument did not undergo statistical validity and reliability testing for several reasons. First, the instrument had already been validated by experts in the field, making empirical validity testing unnecessary. Second, the test produced binary data (correct or incorrect answers), which poses limitations for conventional validation methods. Third, quantitative validity and reliability testing require a large sample size to ensure stable and interpretable results. Calculations such as item-total correlation or reliability tests (e.g., KR-20 or Cronbach's Alpha) would be unreliable with a small sample. Therefore, expert validation was deemed sufficient to ensure the quality of the test.

Qualitative data was collected through a list of open-ended questions given to the students in the form of an interview. This data revealed student perception towards using ChatGPT in learning mathematics. To confirm the validity of the questions, two professional technology education lecturers were asked for their opinions and thoughts on the evaluation and improvement regarding the questions about the aforementioned interview.

Quantitative data were collected through pre-tests and post-tests. The results were subsequently and statistically analyzed using SPSS26. The statistical analyses were descriptive (mean and standard deviation) and inferential (independent t-test). The t-test was utilized to ascertain whether there was a statistically significant difference in mean scores between the experimental and control groups (Mishra et al., 2019).

Qualitative data were collected through student responses to a short essay and then analyzed using thematic analysis. The student responses were systematically coded into themes and categories (Forman & Damschroder, 2007) to identify the advantages and disadvantages of using ChatGPT and explore student perceptions regarding the technology.

Results

Quantitative results

Quantitative data were obtained through pre-tests and post-tests done by students for each experimental and control class. The average test results in the experimental class with a sample of 39 students increased from 36.026 to 40.128. Meanwhile, the control class with a sample of 37 students also increased from 39.054 to 40.540. Analyzing the results with SPSS25 showed the skewness and kurtosis values in both the experimental class (Skewness = 0.945, 0.143; Kurtosis = -0.364, -1.566) and the control class (Skewness = 1.354, 1.094; Kurtosis = 0.031; -0.114) are in the range of -2 to +2. It means that the data is normally distributed and homogeneous. Thus, it ensures that parametric statistical testing can be performed on quantitative data.

Description of student mathematics learning outcomes

The average learning outcomes of students in ChatGPT-integrated and conventional learning only experienced a low increase. In the class with ChatGPT implementation, it was noted that in the pre-test, only 2.56% (N=1) of students exceeded the passing grade (KKM). After the post-test, it increased to 7.79% (N=3) students, with an average increase of 4.102. Meanwhile, in the conventional class, the percentage of students who passed the pre-test and post-test was the same at 8.11% (N=3), with an average increase of 1.486. Descriptively, the increase in the value of the learning outcomes based on pre-test and post-test results between experimental and control classes was relatively slight.

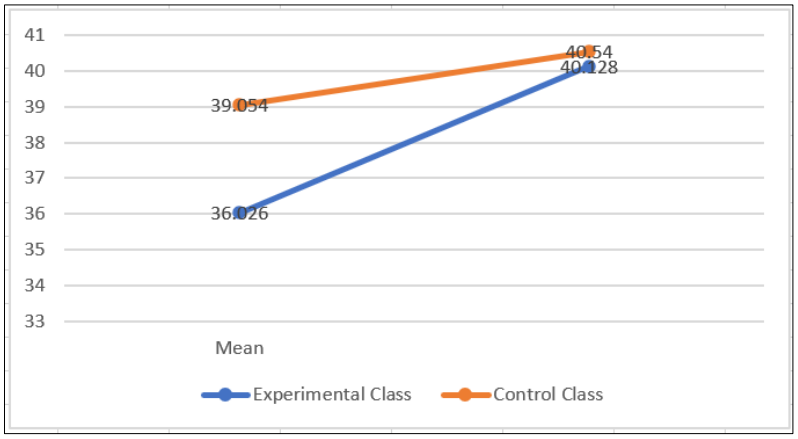

Figure 2. Pretest and posttest results between control and experimental classes

Furthermore, if examined from an effort to explore the use of ChatGPT in learning, specifically in the gender aspect, the average test scores of female (N = 20) and male (N = 19) students in the pre-test were 35.250 and 36.842, respectively. The post-test results revealed a slight increase in each gender category, namely 39.250 and 41.052. From this, there is no significant difference in score improvement between male and female students from using ChatGPT in mathematics learning. It means that neither males nor females were affected by the use of ChatGPT in learning mathematics.

The tests in this study were divided into two types, namely LOTS questions (9 questions) with a maximum score of 55 and HOTS questions (11 questions) with a maximum score of 45.

When viewed from this aspect, with LOTS questions, the average student test score decreased slightly from 23.205 in the pre-test to 21.538 in the post-test. As for HOTS questions, the average student test score increased from 12.820 to 18.590. From these results, it was found that there was an increase in student learning outcomes on HOTS-type questions compared to LOTS-type questions with the use of ChatGPT.

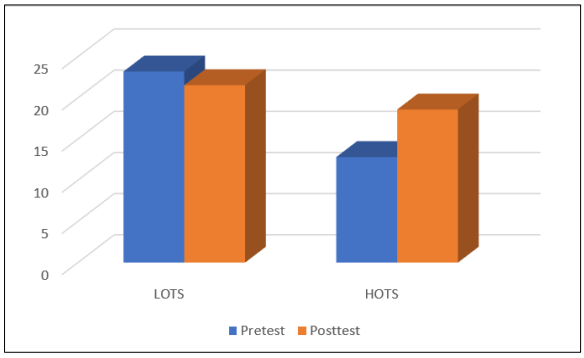

Figure 3. Pretest and Posttest Results of LOTS and HOTS questions for the experimental class

Conventional mathematics learning (control class)

Paired sample t-test was used to test whether there was a difference in students' pre-test and post-test results in the control class. Data on the pre-test (M = 39.054; SD = 16.492) and post-test (40.540; SD = 16.107) in the control class (N = 37) were tested using SPSS which showed there was a significant difference in students' mathematics learning outcomes between the pre-test and post-test of the control class (p-value = 0.026; $p < 0.05$) with a small effect size (d = 0.091). Conventional learning has a negligible effect on improving students' mathematics learning outcomes.

Table 1. SPSS test results comparison of control classes

		Paired Differences					T	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Pretest_Kontrol - Postest_Kontrol	-1.48649	3.88460	.63862	-2.78168	-.19130	-2.328	36	.026

ChatGPT-integrated mathematics learning (experiment class)

Analysis using paired sample t-test was also conducted to test whether there was a difference in students' learning outcomes in the experimental class. The statistical test results showed that there was a significant difference (p-value = 0.013; $p < 0.05$) between the pre-test (M = 36.026; SD = 14.243) and post-test (M = 40.128; 16.563) of students' learning outcomes with the integration of ChatGPT. The effect size of the pre-test and post-test results showed a small effect (0.265). Learning mathematics using ChatGPT integration has a negligible effect on improving students' mathematics learning outcomes.

Table 2. SPSS test results comparison of experiment classes

		Paired Differences					T	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Pretest_Eksperimen - Postest_Eksperimen	-4.10256	9.79203	1.56798	-7.27677	-.92835	-2.616	38	.013

A comparison of conventional mathematics Learning and ChatGPT-integrated learning on student learning outcomes

The comparison of students' learning outcomes between experimental and control groups was analyzed using an independent sample t-test with the post-test scores. The students' mathematics learning outcomes showed no significant difference (p-value = 0.913; $p > 0.05$). It revealed that the scores obtained by students between conventional learning and ChatGPT-integrated learning did not differ. In other words, integrating ChatGPT in mathematics learning does not impact student learning outcomes.

Table 3. SPSS test results comparing control class and experimental class

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
Hasil_ Belajar	Equal variances assumed	.669	.416	-.110	74	.913	-.412	3.751	-7.885	7.061
	Equal variances not assumed			-.110	73.952	.913	-.412	3.748	-7.880	7.055

Qualitative findings

Thematic analysis was used to understand students' views on using ChatGPT integration in mathematics learning. Qualitative data was obtained from short essays completed by 39 students who had previously used ChatGPT while learning. Students were given S codes sequentially, starting from S1, S2, and up to S39. From the students' responses, the following themes emerged.

Efficiency and speed in learning

Students perceived that using ChatGPT in learning maths significantly improved their efficiency and speed in understanding and completing tasks. Many students found it easy to access answers quickly, which allowed them to focus more on the learning process without getting stuck on problems that are difficult or take a long time to answer. ChatGPT provided instant responses, allowing students to make better use of study time and complete assignments more efficiently. In this case, ChatGPT not only accelerates problem-solving but also the process of student understanding. This efficiency contributed directly to improved learning

performance, especially when students need immediate and precise solutions to more challenging problems. Despite some notes regarding the accuracy, the overall speed of ChatGPT was positively welcomed by students as it helped them maximize their study time and improved the overall effectiveness of the learning process. Here are some relevant responses.

> "When I use ChatGPT, I feel very helped and I find it very fast to find answers and I also find it very comforting." (S7)
>
> "My experience when using ChatGPT is that I can find answers using ChatGPT and I can finish exercises fast and ChatGPT can be used with every subject." (S21)

Improvement in understanding and knowledge

Student responses indicated that ChatGPT effectively improved their understanding. Many students reported that ChatGPT helped them understand maths concepts better through clear and detailed explanations. Some students noted that by using ChatGPT, they could easily get additional explanations on difficult material and overcome difficulties in solving maths problems. ChatGPT's ability to provide complete and in-depth answers allowed students to learn more independently and gain a deeper understanding of the material being taught. In addition, students felt that ChatGPT assisted them in learning new mathematical formulas and concepts more effectively, supporting their increased knowledge in the field. Here are some representative student responses.

> "A lot of materials or answers, math can be fun and the questions are easy to answer with ChatGPT." (S13)
>
> "Can generate answers and explanations more detailed and easy to understand by students and can answers math questions whether it's hard or easy." (S35)

Accuracy and quality of answers

Student responses indicated that the results of using ChatGPT integration in learning mathematics varied in accuracy and quality of answers. Many students appreciated ChatGPT's ability to provide quick and generally correct answers, which helped them solve problems and understand the material. However, some students reported situations where ChatGPT answers were not always accurate or in line with the textbook material, which sometimes caused confusion and required additional verification from other sources. Despite some notes of discrepancies, students found ChatGPT helpful overall in providing the necessary answers and explanations. However, they needed to check the accuracy of the answers to ensure they matched the subject matter provided by the teacher.

> "My experience in using ChatGPT is in learning something useful and can provide me useful answers." (S2)
>
> "I find it easy to learn because ChatGPT can search for answers for difficult equations. Can also search for equations for lowest common denominator, greatest common denominator, integer multiplication and others. If ChatGPT is given a problem, it can easily to solve it." (S19).

Support in overcoming learning difficulties

According to student feedback, ChatGPT provided meaningful support in overcoming their learning difficulties. Many students claimed that ChatGPT helped them understand confusing concepts when faced with complex problems. They appreciated ChatGPT's ability to provide clear explanations and relevant solutions, which made it easier for them to understand the material and complete assignments without getting stuck for too long. Some students also felt more confident doing homework and facing maths challenges with help from ChatGPT. While there were a few caveats regarding the accuracy of answers, ChatGPT provided the additional support needed to overcome learning difficulties and improve understanding of the subject matter.

> "When I use ChatGPT to ask questions, turns out ChatGPT can answer questions I find difficult." (S4)
>
> "Very good, can help us in searching for math exercises that are difficult." (S10).

Discussion

The results of this study indicate that the use of ChatGPT in learning mathematics did not affect improving number sense learning outcomes of junior high school students, especially 7th-grade students. This finding clarified previous studies regarding the use of ChatGPT, which did not significantly affect students' learning outcome performance in non-mathematics subjects. For example, the study by Xue et al. (2024) explained that using ChatGPT does not influence students' academic achievement in programming subjects. A logical reason that might explain this is that both mathematics and programming require deep concept understanding and problem-solving skills that ChatGPT'sChatGPT's assistance may not adequately support. ChatGPT may be useful for providing information or clarification, but it is less effective in replacing intensive practice and structured learning (Abdulla et al., 2024).

Furthermore, several factors may explain why learning with ChatGPT did not significantly improve number-sense learning outcomes. Firstly, the fact that ChatGPT was only permitted during maths lessons meant that the frequency of its use was low, resulting in students not having sufficient opportunities to make optimal use of this tool. The policy restricting smartphone use in many schools also limits students' access to ChatGPT, reducing their opportunities to explore mathematical concepts independently. This limitation resulted in sub-optimal learning outcomes, as students cannot actively use technology to support learning outside the classroom. ChatGPT integration regarding student learning outcomes is influenced by the frequency of use factor (Zhou & Li, 2023). In addition, 7th-grade junior high school students are generally still dependent on teacher guidance, making them less independent in learning (Lapan et al., 2001). With ChatGPT relying on self-initiative, this limitation hinders optimal utilization of the technology. Consequently, according to researchers, low learning independence may contribute to ChatGPT's minimal impact on students' mathematics learning outcomes.

In particular, this study showed that students' learning outcomes with HOTS questions improved, while the same did not happen for LOTS questions. It might have been driven by ChatGPT's interactive and adaptive nature, which supports the development of critical and analytical thinking skills. ChatGPT provided step-by-step guidance and in-depth responses that helped students decipher HOTS questions. With dynamic feedback, students were likely to be more encouraged to think critically, creating a richer learning experience than LOTS of questions, which rely more on memorization. The findings also reinforce that integrating AI into learning can stimulate creativity, critical thinking, and problem-solving skills (Celik et al., 2024; Silitonga et al., 2024). However, this finding needs to be further explored since the analysis conducted in this study was only descriptive.

The findings further showed that students confirmed that ChatGPT could facilitate several beneficial aspects of learning mathematics. Firstly, ChatGPT increased efficiency and speed in the learning process. It is consistent with the findings of Cascella et al. (2023) that ChatGPT can synthesize relevant information, making it easier for students to obtain the desired information quickly without spending extended time reading various sources. Students can access knowledge more efficiently, which ultimately accelerates their learning process. Secondly, ChatGPT can deepen students' understanding and knowledge, in line with Ruiz et al. (2023) and Wardat et al. (2023), in which ChatGPT is recognized for improving students' maths skills and supporting educational success. ChatGPT equips students with better foundational knowledge in maths and provides easy access to a wide range of learning topics. ChatGPT enables students to deepen their understanding of the concepts taught by providing easy-to-understand and structured information, directly improving their academic knowledge. Thirdly, the answers generated by ChatGPT are of reliable quality and accuracy and can be connected to the theory expressed by Remoto (2023), which mentions that ChatGPT can respond appropriately to chat requests, especially in problem-solving and mathematical proofs. ChatGPT relies on AI-based algorithms trained on a wide range of data, enabling the system to provide relevant and accurate solutions to questions. Finally, ChatGPT provides support in overcoming learning difficulties by the theory proposed by (Tlili et al., 2023), which states that ChatGPT can help students overcome challenges in learning through the provision of conversational support. ChatGPT's ability to respond interactively to questions and provide easy-to-understand explanations helps students when facing difficulties understanding the material.

The quantitative findings showed that using ChatGPT in learning mathematics did not directly improve 7th-grade students' learning outcomes in number sense learning. The implications of these results indicated that relying on technology such as ChatGPT alone may not be enough to spur significant improvements in academic performance (Gómez Cano & Colala Troya, 2023; Lee & Paul, 2023). It emphasizes the importance of a more comprehensive approach to utilizing educational technology. Educators must recognize that while ChatGPT can be a helpful support tool, its effectiveness depends mainly on integrating more diverse and in-depth teaching methods. Learning mathematics, in particular, demands more intense interaction and an in-depth understanding of concepts, which often requires various

pedagogical approaches (Bashir, 2023). In this context, ChatGPT was more appropriately positioned as a complement to learning rather than a primary solution.

Although quantitatively, the use of ChatGPT in learning mathematics did not significantly improve number sense learning outcomes, the qualitative findings provide a deeper insight into the benefits perceived by students. Students reported that ChatGPT helped improve efficiency and speed in learning, provided better understanding, and offered accurate and quality answers. The implications of these findings suggest that ChatGPT, while not directly improving academic outcomes, has the potential to provide significant support in the cognitive and affective aspects of learning (Almulla & Ali, 2024). Educators can utilize this technology to support students' self-learning, especially when answering questions or simplifying complex concepts. The teacher's role could focus on directing students to use ChatGPT as a tool that complements, rather than replaces, direct interaction with the teacher and learning materials.

This study fills a research gap by explicitly examining the role of ChatGPT in learning number concepts, an area that remains underexplored in mathematics education. Unlike previous studies that broadly assess AI in learning (Alneyadi & Wardat, 2023; Rane, 2023; Remoto, 2024), this research provides empirical evidence of its practical impact through a quasi-experimental approach. The findings contribute to a deeper understanding of how ChatGPT can be integrated with traditional teaching methods to enhance students' conceptual grasp of numbers.

The limitation of this study was the relatively small sample size, with the results not being widely relatable to a larger population. In addition, this study only focused on numbers at the junior high school level, which limits the scope of the findings to specific aspects of mathematics learning. This approach might not have covered a variety of other mathematical concepts, making the results not applicable to different topics or levels of education. Further research with larger samples and a wider range of materials is needed to strengthen and extend these findings.

Conclusion

This study evaluated the effectiveness of using ChatGPT in enhancing 7th-grade students' learning outcomes in number concepts. The quantitative results indicated that ChatGPT did not lead to a significant improvement in students' learning outcomes. However, the qualitative findings revealed that students perceived several advantages, such as increased efficiency in learning, better comprehension of number concepts, improved accuracy in solving problems, and support in overcoming learning difficulties. While ChatGPT did not directly impact academic performance, it was a valuable tool in facilitating the learning process, promoting independent learning, and understanding complex concepts. To optimize learning outcomes in number concepts, it is crucial to integrate ChatGPT with traditional teaching methods. Therefore, educators are encouraged to use ChatGPT as a supplemental tool to enhance students' learning experiences while ensuring the continued importance of teachers' roles and direct classroom interaction.

Acknowledgment

The authors give their most significant appreciation to LP2M IAIN Manado and Junior High School State 10 Manado, who have supported the implementation of the research until it was completed successfully.

Conflicts of Interest

There is no conflict of interest in this research. In addition, the authors have completed the ethical issues, including plagiarism, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies.

Funding Statement

The author thanks LP2M IAIN Manado for its support and funding through Mandatory Research Grant No. P.037/In.25/LP2M/KP.01.2/05/2024.

Author Contributions

Kadir: Conceptualization, writing - original draft, editing, visualization and formal analysis; **Gita F. Jafar**: data analysis and interpretation; **Benjamin B. Mangila**: Methodology, validation and monitoring.

References

Abah, J. (2020). An appeal in the case involving conventional teaching: Emphasizing the transformation to enhanced conventional teaching in mathematics education. *VillageMath Educational Review (VER)*, *1*(1), 1–10. https://doi.org/10.5281/zenodo.3860320

Abdulla, S., Ismail, S., Fawzy, Y., & Elhaj, A. (2024). Using ChatGPT in teaching computer programming and studying its impact on students performance. *Electronic Journal of E-Learning*, *22*(6), 66–81. https://doi.org/10.34190/EJEL.22.6.3380

Almulla, M., & Ali, S. I. (2024). The changing educational landscape for sustainable online experiences: Implications of ChatGPT in Arab Students' learning experience. *International Journal of Learning, Teaching and Educational Research*, *23*(9), 285–306. https://doi.org/10.26803/ijlter.23.9.15

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

Bashir, S. (2023). Pedagogy of Mathematics. *International Journal of Basic Sciences and Applied Computing*, *10*(2), 1–8. https://doi.org/10.35940/ijbsac.b1159.1010223

Bergman, M. M. (2008). *Introduction: Whither mixed methods?* Sage publications.

Campbell, D. T., & Stanley, J. C. (2015). *Experimental and quasi-experimental designs for research*. Ravenio books.

Cascella, M., Montomoli, J., Bellini, V., & Bignami, E. (2023). Evaluating the feasibility of ChatGPT in healthcare: An analysis of multiple clinical and research scenarios. *Journal of Medical Systems*, *47*(1), 1-5. https://doi.org/10.1007/s10916-023-01925-4

Celik, I., Gedrimiene, E., Siklander, S., & Muukkonen, H. (2024). The affordances of artificial

intelligence-based tools for supporting 21st-century skills: A systematic review of empirical research in higher education. *Australasian Journal of Educational Technology*, *SE-Articles*. https://doi.org/10.14742/ajet.9069

Charles, R. I., & Carmel, C. A. (2005). Big ideas and understandings as the foundation for elementary and middle school mathematics. *Journal of Mathematics Education*, *7*(3), 9–24.

Creswell, J. W., Plano Clark, V. L., & Garrett, A. L. (2008). Methodological issues in conducting mixed methods research designs. *Advances in Mixed Methods Research*, *1*(1), 66–83.

Forman, J., & Damschroder, L. (2007). Qualitative content analysis. In *Empirical methods for bioethics: A primer* (pp. 39–62). Emerald Group Publishing Limited.

Frieder, S., Pinchetti, L., Chevalier, A., Griffiths, R.-R., Salvatori, T., Lukasiewicz, T., Petersen, P. C., & Berner, J. (2023). *Mathematical Capabilities of ChatGPT*. http://arxiv.org/abs/2301.13867

Gómez Cano, C. A., & Colala Troya, A. L. (2023). Artificial Intelligence applied to teaching and learning processes. *LatIA*, *1*, 2. https://doi.org/10.62486/latia20232

Goos, M., Vale, C., Stillman, G., Makar, K., Herbert, S., & Geiger, V. (2020). *Teaching secondary school mathematics: Research and practice for the 21st century*. Routledge.

Hemmi, K., Krzywacki, H., & Liljekvist, Y. (2019). Challenging traditional classroom practices: Swedish teachers' interplay with Finnish curriculum materials. *Journal of Curriculum Studies*, *51*(3), 342–361.

Lapan, R. T., Gysbers, N. C., & Petroski, G. F. (2001). Helping seventh graders be safe and successful: A statewide study of the impact of comprehensive guidance and counseling programs. *Journal of Counseling & Development*, *79*(3), 320–330. https://doi.org/10.1002/j.1556-6676.2001.tb01977.x

Lee, J., & Paul, N. (2023). A review of pedagogical approaches for improved engagement and learning outcomes in mathematics. *Journal of Student Research*, *12*(3), 1–9. https://doi.org/10.47611/jsrhs.v12i3.5021

Leelavathi, R., & Surendhranatha, R. C. (2024). ChatGPT in the classroom: navigating the generative AI wave in management education. *Journal of Research in Innovative Teaching & Learning*, *ahead-of-p*(ahead-of-print). https://doi.org/10.1108/JRIT-01-2024-0017

Mishra, P., Singh, U., Pandey, C. M., Mishra, P., & Pandey, G. (2019). Application of student's t-test, analysis of variance, and covariance. *Annals of Cardiac Anaesthesia*, *22*(4), 407.

Rane, N. (2023). Enhancing mathematical capabilities through ChatGPT and similar generative artificial intelligence: roles and challenges in solving Mathematical Problems. *SSRN Electronic Journal*, *December*. https://doi.org/10.2139/ssrn.4603237

Remoto, J. P. (2024). ChatGPT and other AIs: Personal relief and limitations among mathematics-oriented learners. *Environment and Social Psychology*, *9*(1), 1–13. https://doi.org/10.54517/esp.v9i1.1911

Resnick, L. B. (2020). From protoquantities to operators: Building mathematical competence on a foundation of everyday knowledge. In *Analysis of arithmetic for mathematics teaching* (pp. 373–429). Routledge.

Rudolph Jürgen, Tan Samson, Tan, S. (2023). Journal of Applied Learning & Teaching ChatGPT : Bullshit spewer or the end of traditional assessments in higher education ? *Journal of Applied Learning & Teaching*, *6*(1), 242–263. https://doi.org/10.37074/jalt.2023.6.1.9

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 (Switzerland)*, *13*(10).

https://doi.org/10.3390/app13106039

Silitonga, L. M., Wiyaka, Suciati, S., & Prastikawati, E. F. (2024). *The impact of integrating AI Chatbots and microlearning into flipped classrooms: enhancing students' motivation and higher-order thinking skills BT - innovative technologies and learning* (Y.-P. Cheng, M. Pedaste, E. Bardone, & Y.-M. Huang (eds.); pp. 184–193). Springer Nature Switzerland.

Sowder, J. T. (2020). Making sense of numbers in school mathematics. In *Analysis of arithmetic for mathematics teaching* (pp. 1–51). Routledge.

Tang, T., Sha, J., Zhao, Y., Wang, S., Wang, Z., & Shen, S. (2024). Unveiling the efficacy of ChatGPT in evaluating critical thinking skills through peer feedback analysis: Leveraging existing classification criteria. *Thinking Skills and Creativity*, *53*, 101607. https://doi.org/10.1016/j.tsc.2024.101607

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

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

Wijaya, A., van den Heuvel-Panhuizen, M., & Doorman, M. (2015). Opportunity-to-learn context-based tasks provided by mathematics textbooks. *Educational Studies in Mathematics*, *89*, 41–65.

Xue, Y., Chen, H., Bai, G. R., Tairas, R., & Huang, Y. (2024). Does ChatGPT help with introductory programming? an experiment of students using ChatGPT in CS1. *Proceedings - International Conference on Software Engineering*, 331–341. https://doi.org/10.1145/3639474.3640076

Zalima, E. I., Njanji, F. P., Lasmiatik, L., Agustina, L., Dela, M., & Ambarawati, M. (2020). Analysis of students' difficulties in solving word problems on arithmetic operations with mixed Fractions [in Bahasa]. *Prismatika: Jurnal Pendidikan dan Riset Matematika*, *2*(2), 46–54.

Zhou, L., & Li, J. (2023). The impact of ChatGPT on learning motivation: A study based on self-determination theory. *Education Science and Management*, *1*(1), 19–29. https://doi.org/10.56578/esm010103