



ANALYSIS OF DIFFERENCES IN THE RESULTS OF MATHEMATICS COGNITIVE TEST OF GRADE XI STUDENTS USING CHAT-GPT AND OPEN BOOK AT SMAN 7 KEDIRI

Almas Ardhia Pramisti¹, Faradila Ramadhani^{2*}, Arafah Husna³, Ence Surahman⁴

^{1,2,3,4}State University of Malang, Jl. Semarang No. 5 Malang, East Java, Indonesia

*Author of correspondence, Email: faradila.ramadhani.2201216@students.um.ac.id

Abstract

This study aims to determine the effect of using ChatGPT as an aid in solving mathematics problems on the cognitive test results of eleventh-grade students on Trigonometry material at SMAN 7 Kediri City. The study used a quasi-experimental method with a static non-equivalent group design. The study sample consisted of two classes, namely the control class (using the open book method) and the test class (using ChatGPT). Data were obtained through cognitive tests, observations, and questionnaires. The results of the Mann-Whitney test showed a significant difference between the two classes ($p < 0.05$), with the average score of the test class being higher (98.57) than the control class (58.50). Effectiveness analysis using Cohen's d showed a large effect ($d = 2.14$). Qualitative results also supported this finding, where students showed positive perceptions of the use of ChatGPT. Thus, ChatGPT was proven to have a positive effect on students' cognitive mathematics learning outcomes.

Keywords: ChatGPT, Cognitive Test, Mathematics, Trigonometry

1. Introduction

The rapid development of the digital era has made artificial intelligence (AI) an integral part of various aspects of life, including education. One AI innovation that has attracted significant attention in recent years is ChatGPT, a machine learning-based language model developed by OpenAI. Designed to understand and generate text with a natural language structure resembling human conversation, the latest version, ChatGPT-4, shows significant improvements in speed, accuracy, and contextuality compared to previous (Islam & Moushi, 2024). These findings are supported by research by (L. Halimah & Jamilah, 2024), which shows that integrating AI into learning can improve learning outcomes and student engagement.

The digital transformation in education has opened up significant opportunities to transform teaching and evaluation methods. Technologies like ChatGPT enable more adaptive, personalized, and immersive learning because students can interact directly with systems capable of providing instant feedback and diverse explanations (Nguyen et al., 2023). Thus, education is no longer tied to traditional face-to-face and memorization-based models, but rather is evolving toward exploration-based and problem-solving learning, where students play an active role in constructing their own knowledge.

In the context of mathematics learning, the main challenges faced by students often relate to difficulties in understanding abstract concepts and applying formulas in problem-solving. Mathematics, known as the "queen of sciences" (Afsari et al., 2021), plays a crucial role in developing logical, systematic, and critical thinking skills. However, according to the PISA

report (OECD, 2024), Indonesian students' mathematics learning outcomes are still below the global average. Research by (Woo et al., 2024) shows that students with better prompt engineering skills—namely, the ability to effectively ask questions to AI—tend to achieve more optimal learning outcomes.

This emphasizes the importance of adopting a learning approach relevant to the characteristics of the 21st century. The use of ChatGPT in mathematics learning offers a potential solution to these problems. ChatGPT not only provides step-by-step solutions to mathematics problems but also comprehensively explains the underlying concepts. (Yuntaro & Yudhanegara, 2024) reported that the use of ChatGPT improved mathematics problem-solving skills at SMA Negeri 2 Bantaeng. Meanwhile, (Masluha & Zainudin, 2024) showed that the flipped classroom model supported by ChatGPT significantly improved mathematics learning outcomes for students at MTs Darul Ulum.

However, technology isn't the only component that determines learning effectiveness. Evaluation methods also play a crucial role in achieving learning outcomes. One increasingly used evaluation approach is the open-book method, which allows students to access the material during exams. (Lynd-Balta, 2020) found that this method can reduce anxiety and improve conceptual understanding, as students focus more on analyzing and synthesizing information rather than simply memorizing. This approach is particularly relevant in mathematics learning, which demands a deep understanding of complex concepts.

Trigonometry is a mathematics topic often considered difficult by high school students. Concepts such as special angles, trigonometric identities, inverse functions, and their application in real-world contexts require a strong understanding and analytical thinking skills. According to (Harjono et al., 2020), more than 70% of students experience difficulty understanding trigonometry, particularly in applying formulas and solving word problems. This difficulty is exacerbated by students' poor ability to transfer knowledge from theory to practice. This is where technologies like ChatGPT have the potential to be of great benefit. (Lo, 2023) states that ChatGPT can function as a virtual tutor, helping students think computationally and understand complex material gradually.

However, concerns remain regarding the potential negative impacts of reliance on AI. (Auna & Hamzah, 2024) warn that excessive use of AI can diminish students' initiative and critical thinking skills. Therefore, a constructivist-based approach is crucial, in which students actively construct knowledge through their own experience (Fosnot, 2013). In this context, ChatGPT should function as a facilitator of intellectual exploration, not simply a tool for answering questions. Students can learn by analyzing AI responses, evaluating the logic of solutions, and developing follow-up questions—an approach consistent with active learning (Güner & Er, 2025)

However, in practice, the implementation of learning technology still faces challenges. At SMAN 7 Kediri City, for example, interviews with mathematics teachers revealed that students' grades in the subject tend to be low, and none have achieved academic success in mathematics. Teachers are also unfamiliar with using AI in the learning process. This situation reflects a gap between the technology's potential and the readiness for implementation in schools.

(Bekdemir, 2024) emphasized that the success of AI integration in learning depends heavily on students' digital literacy and clear pedagogical guidance from teachers. Therefore, this study aims to empirically evaluate how different learning methods—namely the use of ChatGPT as an aid compared to the open-book method—affect students' cognitive test results. This aligns with the analysis of (Parmiti et al., 2021) who stated that the effectiveness of the open-book method is strongly influenced by students' ability to navigate and apply information, not simply access it. This study seeks to answer whether the use of AI technology such as ChatGPT can significantly impact students' cognitive achievement in Trigonometry. Unlike previous research that focused more on students' motivation or perceptions of ChatGPT, this study specifically examines the direct effect of its use on Mathematics cognitive test results, by comparing two approaches: the use of AI technology and the conventional open-book method. The focus on Trigonometry provides room for a more in-depth analysis of the effectiveness of each approach.

2. Method

This study employed a quantitative approach with a quasi-experimental design, specifically a static-group comparison design, which is particularly suitable for use in pre-formed classrooms without randomization. This design allows researchers to compare two groups with different treatments—a test class using ChatGPT and a control class using the open-book method—to examine differences in students' cognitive learning outcomes after a single trigonometry practice session. A static-group comparison design is suitable when pure experimental conditions are not feasible and internal validity can be controlled through matching subject characteristics.

This research was conducted at SMAN 7 Kediri City during April 2025, with the research subjects being grade XI students majoring in Mathematics and Natural Sciences who were divided into two classes, namely class XI-1 as a test class that used ChatGPT assistance in solving math problems, and class XI-2 as a control class that worked on problems using the open book method without the assistance of AI technology. The number of research samples consisted of 28 students in the test class and 30 students in the control class, so that the total sample was 58 students. The sampling technique was carried out by purposive sampling, with the consideration that both classes have equivalent characteristics in terms of curriculum, initial competencies, and learning facilities. This equivalence is assumed based on information from the teaching teacher and the homogeneity of the learning context applied in the school.

The main instrument in this study was 30 multiple-choice questions related to trigonometry. After validity and reliability testing, 15 questions were selected for use in hypothesis testing. The question grid was systematically designed to cover aspects of conceptual understanding, application of basic trigonometric identities, real-life applications, function graphs, period and amplitude, inverse functions, additional identities, and contextual story problems. Expert validation showed that 92.3% of the questions met standards (Widarti et al., 2022)

Table 1. Mathematics Cognitive Test Question Grid

No	Aspect	Indicator	Sub-Indicators	Number of Questions	Question Number
1	Concept Understanding	Understanding the definition of trigonometric functions	Sin, cos, tan in triangles; special angles	6	1–6
2	Application of Formula	Using basic trigonometric identities	Basic trigonometric identities	4	7–10
3	Life Application	Solving contextual problems	Angles of elevation and depression	4	11–14
4	Function Graph	Understanding trigonometric function graphs	$y = \sin x, \cos x, \tan x$	4	15–18
5	Period & Amplitude	Determining the period and amplitude	$y = a \sin (bx)$	4	19–22
6	Inverse Function	Understanding inverse trigonometric functions	arcsin, arccos, arctan	3	23–25
7	Additional Identity	Using double angle and half angle identities	$\sin 2x, \cos 2x, \tan 2x$	3	26–28

This instrument was declared valid based on a validity test using Pearson correlation, with 26 of the 30 questions having a calculated r value > 0.2146 (r table at $\alpha = 0.05$). Questions P10, P23, P29, and P30 were invalid, but P30 was still used because it represented an important indicator and was close to the validity limit. The reliability test using Cronbach's Alpha produced a value of $\alpha = 0.835$, indicating high internal consistency (Arikunto, 2016; Cohen et al., 2018).

In addition to the test instruments, the researchers also used different observation protocols for each class to observe student interactions while solving problems. Observations covered aspects of learning focus, collaboration, answering strategies, independence, and completion time. Observations were conducted by the researchers as direct observers, and the results were analyzed descriptively to complement the quantitative data. For the test class, observations focused on the use of ChatGPT, how students constructed prompts, understood answers, and evaluated AI solutions. Meanwhile, in the control class, observations covered how students searched for information from textbooks and interacted with their notes (Braun & Clarke, 2006).

This study also used a student perception questionnaire consisting of 25 closed-ended statements based on a Likert scale (1–5), covering aspects of learning interest, motivation, effectiveness, attitudes toward technology, and learning outcomes. This questionnaire was only administered to the test class after the test was administered to determine the extent to which students assessed the use of ChatGPT in helping them understand and solve mathematics problems (Zein et al., 2024)

The research implementation process was divided into four main stages: (1) Preparation, including the preparation of questions in Google Form, instrument validation, and preparation of observation protocols and questionnaires; (2) Implementation of the posttest which was carried out simultaneously in two classes with different conditions according to each treatment; (3) Observation of students while working on the questions, and (4) filling out the questionnaire by the test class. The research was conducted in one meeting to ensure environmental control remained stable and free from external factors.

To complement the quantitative analysis, qualitative data from observations and questionnaires were analyzed using thematic analysis according to (Braun & Clarke, 2006), which involves data familiarization, initial coding, theme identification, theme review, and final interpretation. This analysis aimed to gain a deeper understanding of how students interpreted the use of ChatGPT as a learning tool and how their experiences differed from those of the open-book method. Throughout the research process, the researchers upheld research ethics by seeking permission from the school, explaining the purpose and benefits of the study to students, and maintaining the confidentiality of participants' identities. All data collected was used solely for academic purposes.

3. Results and Discussion

3.1 Result

This study aims to analyze the differences in students' cognitive mathematics test results between the group using ChatGPT (experimental class) and the group using open book (control class) on trigonometry. Data were collected from test results given to both classes after practice sessions. The following is a quantitative description of the research results.

Table 2. Results of the Test of Question Item Validity

Question Item Status	Number of Questions	Question Number
Valid	26	1-9, 11-28
Invalid (Drop)	3	10, 23, 29
Invalid (Used)	1	30

Based on Table 2, it can be concluded that most of the questions developed are valid and can be used to measure students' cognitive abilities accurately.

Furthermore, a reliability test using Cronbach's Alpha yielded a value of 0.835. This value is considered reliable because it is >0.70 , indicating excellent internal consistency of the instrument.

Table 3. Results of the Reliability Test of Question Items

Cronbach's Alpha	N of Item
.835	30

Based on Table 3, the test instrument used is classified as very reliable so it is suitable for use in research.

Test Results of Question Difficulty Level

A difficulty level test was conducted to determine the extent to which the questions differentiated students' abilities. The analysis showed that there were 8 questions categorized as very easy, 15 questions as easy, 5 questions as medium, 1 question as difficult, and no questions as very difficult. Questions categorized as very easy were not used in the hypothesis test because they were deemed less representative in assessing students' overall abilities.

Table 4. Distribution of Question Difficulty Levels

Difficulty Category	Number of Questions
Very Easy	8
Easy	15
Currently	5
Difficult	1
Very Difficult	0

Based on Table 4, it can be seen that the majority of questions fall into the easy and moderate categories. Questions in the very easy category were not used in the hypothesis test because they were deemed less representative in optimally measuring student ability.

Descriptive Analysis Results

Descriptive analysis was conducted to determine the average and distribution of posttest scores of students in both groups. The test class (using ChatGPT) obtained an average score of 98.57 with a standard deviation of 3.29. Meanwhile, the control class (using the open book method) obtained an average score of 58.50 with a standard deviation of 26.26. The high average score and low deviation in the test class indicate a better level of consistency in learning outcomes compared to the control class.

Table 5. Descriptive Statistics of Cognitive Test Results

	N	Minimum	Maximum	Mean	Std. Deviation
Test	28	85.00	100.00	98.5714	3.29341
Control	30	10.00	100.00	58.5000	26.26621
Valid N (listwise)	28				

Based on Table 5, the average value of the test class is much higher than the control class, while the low standard deviation indicates that the results of students in the test class are more even and consistent.

Results of Normality and Homogeneity Tests

Based on Table 6, the normality test was conducted using the Shapiro-Wilk test and showed that the data in both groups were not normally distributed (p-value < 0.05). Based on Table 7, the homogeneity of variance test was conducted using Levene's Test and showed that the data were not homogeneous (p-value < 0.05).

Table 6. Normality Test Results

Shapiro-Wilk			
Class	Statistic	df	Sig.
Test	.491	28	.000
Control	.929	30	.047

Table 7. Results of Homogeneity Test

Levene Statistic	df1	df2	Sig
102.014	1	56	.000

Before conducting the hypothesis test, the researchers conducted normality and homogeneity tests. The Shapiro-Wilk test for normality showed that the data were not

normally distributed ($p < 0.05$). Therefore, the non-parametric Mann-Whitney U test was used to compare scores between the two groups.

Table 8. Mann-Whitney U Test Results

	Results
Mann-Whitney-U	21.500
Wilcoxon W	486.500
WITH	-6.409
Aymp. Sig. (2-tailed)	.000

Based on Table 8, the results of the Mann-Whitney U test show a significance value of 0.000 ($p < 0.05$) indicating that there is a significant difference between the learning outcomes of students who use ChatGPT and those who use the open book method..

Cohen's d Analysis Results

To measure the strength of the treatment effect, an effect size calculation was performed using Cohen's d. The mean value for the test class was 98.57 and the mean for the control class was 58.50, with a combined standard deviation of 18.71. The calculation results showed that Cohen's d was 2.14, which is classified as a very large effect.

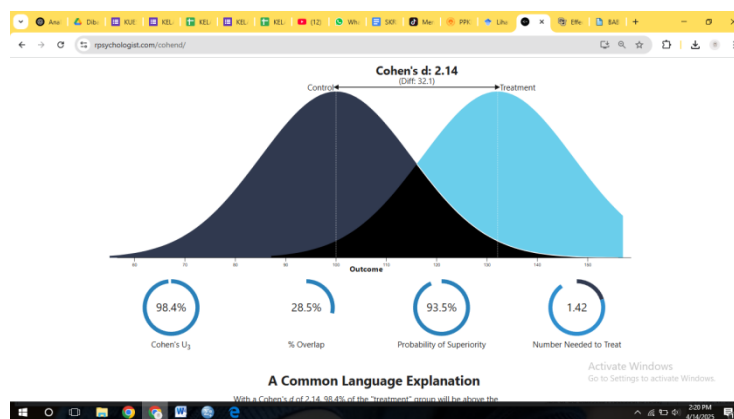


Figure 1. Cohen's D Analysis

Based on the research results, the average score for the test class was 98.57 ($SD = 3.29$) and the control class was 58.50 ($SD = 26.26$). Based on these data, Cohen's d value was 2.14, which according to (Cohen et al., 2017) is categorized as a very large effect. These results indicate that the use of ChatGPT has a significant and practically meaningful impact on improving students' cognitive mathematics test results, particularly in trigonometry.

3.2 Discussion

The findings of this study indicate that the use of ChatGPT significantly improved students' cognitive mathematics test results on trigonometry compared to the open-book method. The test class' average score of 98.57 with a low standard deviation indicated that almost all students successfully completed the problems well and consistently. In contrast, the control class showed lower performance with an uneven distribution of scores, reflecting significant variation in students' ability to understand problems based solely on conventional references. This is in line with research by (Yuntaro & Yudhanegara, 2024) and (Masluha & Zainudin, 2024) which stated that the use of AI technology such as ChatGPT can significantly strengthen students' understanding in solving mathematical problems. ChatGPT functions as a virtual tutor that provides systematic explanations, answers students' questions directly, and enriches their understanding of abstract concepts such as trigonometric functions, graphs, and amplitude calculations. These findings align with (Nugraha & Harsono, 2024) in JKTP, which shows that integrating AI-based technology in creating learning media can significantly improve students' understanding, especially in abstract and complex topics.

Observations showed that students in the test class demonstrated more active and independent learning behavior, such as asking complex questions to ChatGPT, evaluating the answers, and comparing them with their own understanding. In contrast, students in the control class frequently relied on peer discussions and experienced confusion when searching for references in textbooks, particularly when faced with trigonometry application problems. A perception questionnaire distributed to the test class corroborated these findings. The majority of students considered ChatGPT an effective, fast, and engaging tool, especially when dealing with problems requiring in-depth understanding. This reinforces the notion that the use of AI technology not only supports academic performance but also increases student motivation and interest in learning (Ma'amor et al., 2024). (N. Halimah et al., 2024) also emphasized that digital technology-based learning systems, such as LMSs or interactive media, encourage students' self-regulation and active participation in learning.

However, it is important to note that the effectiveness of ChatGPT usage depends on the context and teacher supervision. Some students observed appeared to rely too heavily on the system's answers without critically evaluating them. This demonstrates the importance of digital literacy and ethical technology use in learning, ensuring that students are not merely passive users but also develop critical and reflective thinking skills (Lo, 2023; Rahayu et al., 2022; Subiyantoro et al., 2023) also emphasize the importance of a balanced pedagogical approach between technology and teacher involvement to ensure meaningful understanding for students. Methodologically, although the results show strong effects, this study has limitations because it did not use a pretest to measure initial abilities, so potential bias cannot be completely avoided. Therefore, further studies are recommended with a pretest-posttest design and a longer timescale to comprehensively measure knowledge retention and transfer.

4. Conclusion

Based on the research results and data analysis that have been conducted, it can be concluded that the use of ChatGPT as an aid in solving trigonometry problems significantly improves the results of the Mathematics cognitive test of grade XI students at SMAN 7 Kediri

City. This is indicated by the very striking difference in average scores between the test class and the control class, namely 98.57 compared to 58.50, as well as the results of the Mann-Whitney U test which shows significance at $p < 0.05$. In addition, the Cohen's d value of 2.14 indicates that the use of ChatGPT has a very large effect on student learning outcomes in practice. This finding is in line with the results of previous studies that confirm that the integration of AI technology in learning can support the understanding of complex concepts, increase learning motivation, and provide a more personalized and adaptive learning experience (Masluha & Zainudin, 2024; Wardat et al., 2023; Yuntaro & Yudhanegara, 2024). Through observations and questionnaires, it was found that students also had a positive perception of the use of ChatGPT, which was seen as a tool that could facilitate the process of thinking and solving mathematical problems.

However, despite its significant benefits, ChatGPT cannot be used as a sole benchmark for assessing student abilities and competencies. Teacher mentoring, strengthening digital literacy, and developing critical and reflective thinking skills are necessary to prevent students from passively relying on technology. Furthermore, the limitations of the research design, which lacked a pretest and the short observation period, warrant further research. Overall, ChatGPT has proven to have significant potential as an innovative learning medium in mathematics, particularly in abstract topics like trigonometry. However, its use must be directed, ethical, and integrated with active learning strategies to create a holistic and meaningful learning process.

References

- Afsari, S., Safitri, I., Harahap, S. K., & Munthe, L. S. (2021). Systematic literature review: Efektivitas pendekatan pendidikan matematika realistik pada pembelajaran matematika. *Indonesian Journal of Intellectual Publication*, 1(3), 189–197.
- Auna, H. S. A., & Hamzah, N. (2024). Studi perspektif siswa terhadap efektivitas pembelajaran matematika dengan penerapan chatgpt. *HINEF: Jurnal Rumpun Ilmu Pendidikan*, 3(1), 13–25.
- Bekdemir, Y. (2024). The Urgency of AI Integration in Teacher Training: Shaping the Future of Education. *Journal of Research in Didactical Sciences*, 3(1), 3. <https://doi.org/10.51853/jorids/15485>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Cohen, L., Manion, L., & Morrison, K. (2017). *Research Methods in Education* (8th ed.). Routledge. <https://doi.org/10.4324/9781315456539>
- Fosnot, C. T. (2013). *Constructivism: Theory, perspectives, and practice*. Teachers College Press. [https://books.google.com/books?hl=en&lr=&id=-pIbAgAAQBAJ&oi=fnd&pg=PA1926&dq=Fosnot,+C.+T.+\(2013\).+Constructivism:+Theory,+perspectives,+and+practice.+Teachers+College+Press.](https://books.google.com/books?hl=en&lr=&id=-pIbAgAAQBAJ&oi=fnd&pg=PA1926&dq=Fosnot,+C.+T.+(2013).+Constructivism:+Theory,+perspectives,+and+practice.+Teachers+College+Press.)

- Güner, H., & Er, E. (2025). AI in the classroom: Exploring students' interaction with ChatGPT in programming learning. *Education and Information Technologies*, 30(9), 12681–12707. <https://doi.org/10.1007/s10639-025-13337-7>
- Halimah, L., & jamilah, weti sri nuryani. (2024). INNOVATIVE LEARNING STRATEGIES FOR SOCIAL STUDIES IN THE ERA OF ARTIFICIAL INTELLIGENCE. *International Conference on Aplied Social Sciences in Education*, 1(1), 181–187. <https://doi.org/10.31316/icasse.v1i1.6842>
- Halimah, N., Degeng, M. D. K., & Soepriyanto, Y. (2024). Peranan Tingkat Penerimaan Learning Management System Terhadap Kemampuan Self Regulated Learning. *JKTP: Jurnal Kajian Teknologi Pendidikan*, 7(3), 118–128.
- Harjono, A., gunawan, gunawan, Nisyah, M., Kusdiastuti, M., & Herayanti, L. (2020). Improving Students' Problem-Solving Skills Using Inquiry Learning Model Combined with Advance Organizer. *International Journal of Instruction*, 13(4), 427–442. <https://doi.org/10.29333/iji.2020.13427a>
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410.
- Lynd-Balta, E. (2020). Using Literature and Innovative Assessments to Ignite Interest and Cultivate Critical Thinking Skills in an Undergraduate Neuroscience Course. *CBE—Life Sciences Education*, 5(2), 167–174. <https://doi.org/10.1187/cbe.05-08-0108>
- Ma'amor, H., Achim, N., Ahmad, N. L., Roszaman, N. S., Kamarul Anuar, N. N., Khairul Azwa, N. C. A., Abd Rahman, S. N., & Aqilah Hamjah, N. A. (2024). The Effect of Artificial Intelligence (AI) on Students' Learning. *Information Management and Business Review*, 16(3S(I)a), 856–867. [https://doi.org/10.22610/imbr.v16i3S\(I\)a.4178](https://doi.org/10.22610/imbr.v16i3S(I)a.4178)
- Masluha, M., & Zainudin, Z. (2024). Pengaruh Penerapan Model Pembelajaran Flipped Classroom Berbantuan Chat GPT Terhadap Hasil Belajar Matematika Siswa di MTS Darul Ulum. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, 8(3), 2149–2157. <https://doi.org/10.31004/cendekia.v8i3.3294>
- Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B.-P. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies*, 28(4), 4221–4241. <https://doi.org/10.1007/s10639-022-11316-w>
- Nugraha, C. A., & Harsono, R. J. (2024). Chemistry Storytelling with AI: A Descriptive Research of Ebook Creation for Grade 10 Students. *JKTP: Jurnal Kajian Teknologi Pendidikan*, 7(3), 149–160.
- OECD. (2024). *Transforming education in Indonesia: Examining the landscape of current reforms* (OECD Education Policy Perspectives 88; OECD Education Policy Perspectives, Vol. 88). <https://doi.org/10.1787/9ff8d407-en>

- Parmiti, D. P., Rediani, N. N., Antara, I. G. W. S., & Jayadiningrat, M. G. (2021). The Effectiveness of Local Culture-Integrated Science Learning through Project-Based Assessment on Scientific Attitudes and Science Process Skills of Elementary School Students. *Jurnal Pendidikan IPA Indonesia*, 10(3), 439–446. <https://doi.org/10.15294/jpii.v10i3.31301>
- Rahayu, A. P., Hidayati, D., Susilowati, S., & Jannah, F. B. F. (2022). Pembelajaran Interaktif Melalui Hybrid Learning Dengan Metode Demonstrasi. *JKTP: Jurnal Kajian Teknologi Pendidikan*, 5(3), 326–336.
- Subiyantoro, S., Degeng, I. N. S., Kuswandi, D., & Ulfa, S. (2023). Exploring the impact of ai-powered chatbots (chat gpt) on education: A qualitative study on benefits and drawbacks. *Jurnal Pekommas*, 8(2), 157–168.
- 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.
- Widarti, H. R., Hakim, M. I., & Rokhim, D. A. (2022). The Development of a Virtual Laboratory on Qualitative Chemical Practicum Analysis. *Jurnal Ilmiah Peuradeun*, 10(3), 783. <https://doi.org/10.26811/peuradeun.v10i3.760>
- Woo, D. J., Deliang Wang, Yung, T., & Guo, K. (2024). *Effects of a Prompt Engineering Intervention on Undergraduate Students' AI Self-Efficacy, AI Knowledge, and Prompt Engineering Ability: A Mixed Methods Study*. <https://doi.org/10.13140/RG.2.2.32577.36968>
- Yuntaro, I. S., & Yudhanegara, M. R. (2024). Analisis Klaster Pengaruh Hasil Belajar Matematika dan Bahasa Indonesia Terhadap Pengerjaan Soal Matematika Dibantu ChatGPT. *Didactical Mathematics*, 6(2), 141–150.
- Zein, S., Yaniawati, P., & Mudrikah, A. (2024). The Role and Evaluation of ChatGPT as a Virtual Tutor in Improving Students' Creative and Critical Abilities Reviewed from Probing-Prompting Abilities. *Kreano, Jurnal Matematika Kreatif-Inovatif*, 15(2), 501–517. <https://doi.org/10.15294/a2g5x690>