



Generative AI as Instructional Scaffolding: Effects on Critical Thinking and Classroom Engagement in Secondary Education

Herman¹, Muh. Nasir¹, Anggar Putra¹, Muh. Eka Hidayatullah¹

¹Faculty of Teacher Training and Education, Ngguusuwaru University, Indonesia

*Corresponding Author: Herman

E-mail: nabillaherman06@gmail.com



Article Info

Article history:

Received 5 May 2026

Received in revised form 5 August 2026

Accepted 29 August 2026

Keywords:

Generative Artificial

Intelligence

Instructional Scaffolding

Critical Thinking

Classroom Engagement

Abstract

This study examined the effects of generative artificial intelligence as instructional scaffolding on secondary school students' critical thinking and classroom engagement. A quantitative quasi-experimental method with a nonequivalent pretest-posttest control group design was applied to 64 eighth-grade students at SMP Negeri 1 Kota Bima, comprising 32 students in the experimental group and 32 students in the control group. The experimental group participated in eight generative artificial intelligence-supported learning sessions involving problem identification, information exploration, response verification, collaborative discussion, problem-solving, and reflection, while the control group received conventional instruction. Data were collected through a 20-item critical thinking test, structured classroom observations, and documentation. The data were analyzed using descriptive statistics, normality and homogeneity tests, normalized gain analysis, independent-samples testing, and effect-size analysis. The experimental group achieved a higher posttest mean score than the control group, with scores of 93.75 and 78.30, respectively. The normalized gain was 0.57 in the experimental group and 0.22 in the control group. The differences in posttest and normalized-gain scores were statistically significant, with a large effect size of 0.87. Classroom engagement also reached 92.11 percent in the experimental group compared with 71.80 percent in the control group. These findings demonstrate that generative artificial intelligence effectively supports critical thinking and classroom engagement when used as guided instructional scaffolding rather than as a provider of final answers.

Introduction

The rapid development of generative Artificial Intelligence (AI) has created new opportunities for improving teaching and learning. Generative AI can provide alternative explanations, immediate feedback, learning recommendations, and support for information exploration (Holmes et al., 2019; Ikram et al., 2026; Nacheva, 2024). These capabilities enable AI to function not merely as a source of information but as instructional scaffolding that assists students in completing complex learning tasks while preserving the teacher's role in guiding, monitoring, and evaluating learning (Moneus, Ahmed Mohammed, 2024; UNESCO, 2021; Votto et al., 2021).

The use of generative AI is relevant to the demands of twenty-first-century education, particularly the development of critical thinking, problem-solving, collaboration, digital literacy, and active learning. These competencies are difficult to develop through teacher-centered instruction that emphasizes memorization and the reproduction of information (Ikram et al., 2026). Students need learning experiences that encourage them to investigate problems, evaluate evidence, compare perspectives, construct arguments, and reflect on their conclusions (Aprianto et al., 2025; Tan et al., 2025).

However, the availability of digital technology has not always been followed by meaningful changes in classroom practice (Bagdonaite & Degiene, 2025). Many secondary-school students remain passive during learning and have limited opportunities to analyze information, participate in discussion, and solve problems collaboratively (Sitepu et al., 2026). This condition may weaken their ability to interpret information, evaluate evidence, draw logical conclusions, and communicate reasoned explanations (Zablotska et al., 2026). The problem becomes increasingly important because generative AI may produce inaccurate, incomplete, or contextually inappropriate responses (Saar et al., 2025). Therefore, students need guidance to use AI critically rather than accept its outputs without verification.

Critical thinking involves interpretation, analysis, evaluation, inference, and explanation. These dimensions enable students to assess the credibility of information and formulate conclusions supported by evidence (Georgopoulou et al., 2025). Classroom engagement is also essential because critical thinking develops through active participation, discussion, sustained intellectual effort, problem-solving, and reflection (Agol et al., 2024). Behavioral, emotional, and cognitive engagement therefore provide important conditions for meaningful learning (Wardani & Fiorintina, 2023).

Previous studies have shown that AI-supported learning can improve personalized learning, academic achievement, motivation, problem-solving, and higher-order thinking (Ayanwale & Omeh, 2026). AI can provide adaptive assistance and immediate feedback that encourage inquiry and reflection (Kasneci et al., 2023). Nevertheless, existing studies have more frequently examined AI as a technological tool or focused on higher education (Holmes et al., 2019). Research that positions generative AI as structured instructional scaffolding and simultaneously examines critical thinking and classroom engagement among secondary-school students remains limited (Supriadi et al., 2024). The development of previous research and the position of the present study are illustrated in Figure 1.

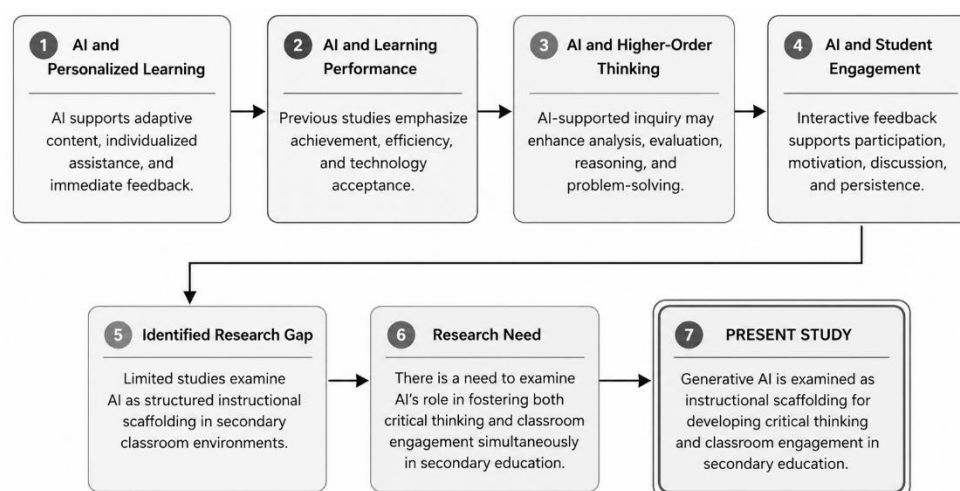


Figure 1. State of the Art of Generative AI, Critical Thinking, and Classroom Engagement (Source: Developed by the authors (2025)).

As illustrated in Figure 1, previous research has primarily examined the contribution of AI to personalized learning, academic performance, higher-order thinking, and student engagement (Tan et al., 2025). However, these areas have generally been investigated separately, while limited attention has been directed toward the pedagogical processes through which AI-supported questioning, feedback, information verification, discussion, and reflection can simultaneously strengthen students' critical thinking and classroom engagement (Gabelaia,

2025). Evidence concerning the use of generative AI as structured instructional scaffolding in secondary education, particularly within the Indonesian context, also remains limited (Tashtoush et al., 2025).

This research gap indicates the need for an instructional design that positions generative AI not as a provider of final answers but as temporary cognitive support. The proposed learning process integrates problem identification, guided information exploration, verification of AI-generated content, collaborative argumentation, problem-solving, and reflection. Through these activities, students remain responsible for evaluating information, constructing arguments, and formulating conclusions. This design is expected to reduce dependence on automated responses and encourage students to use generative AI as a stimulus for reasoning and active participation.

The study is urgent because the rapid use of generative AI among secondary-school students has not always been accompanied by adequate AI literacy and pedagogical guidance. Without structured instructional activities, AI may promote superficial understanding, uncritical acceptance of information, and technological dependence. Conversely, when AI is used through appropriate scaffolding, it can help students identify misconceptions, examine evidence, compare alternative viewpoints, improve arguments, and revise conclusions.

The novelty of this study lies in positioning generative AI as instructional scaffolding rather than merely as educational technology. It also examines critical thinking and classroom engagement as interconnected learning outcomes, analyzes critical thinking through the dimensions of interpretation, analysis, evaluation, inference, and explanation, and provides quasi-experimental evidence from an Indonesian secondary-school context.

Accordingly, this study aims to examine the effects of generative AI-supported instructional scaffolding on secondary-school students' critical thinking and classroom engagement. Specifically, it investigates whether students receiving generative AI-supported instruction achieve higher critical thinking performance and demonstrate stronger active participation, classroom discussion, analytical involvement, problem-solving activity, and learning motivation than students receiving conventional instruction. Based on the theoretical framework and previous findings, it is hypothesized that generative AI-supported instruction significantly improves students' critical thinking and is accompanied by higher behavioral, emotional, and cognitive classroom engagement.

Methods

Research Design and Participants

This study employed a quantitative approach using a quasi-experimental method to examine the effects of generative Artificial Intelligence as instructional scaffolding on secondary school students' critical thinking and classroom engagement. A nonequivalent pretest–posttest control group design was selected because the participants came from existing classes and individual random assignment was not feasible. The experimental group received generative AI-supported instruction, while the control group received conventional teacher-centered instruction. Both groups completed critical thinking pretests and posttests, while classroom engagement was assessed through structured observations during the learning process.

Table 1. Nonequivalent Pretest–Posttest Control Group Design

Group	Pretest	Treatment	Posttest	Classroom Observation
Experimental group	O ₁	X	O ₂	O ₃
Control group	O ₁	C	O ₂	O ₃

Notes: O₁ represents the critical thinking pretest, O₂ represents the critical thinking posttest, O₃ represents classroom engagement observation, X represents generative AI-supported instructional scaffolding, and C represents conventional teacher-centered instruction.

The study was conducted at SMP Negeri 1 Kota Bima during the second semester of the 2025/2026 academic year. The population consisted of all eighth-grade students enrolled at the school. The sample was selected using purposive sampling by considering similarities in academic achievement, classroom characteristics, learning conditions, and recommendations from subject teachers and school administrators.

A total of 64 students participated in the study. The experimental group consisted of 32 students, while the control group consisted of 32 students. Both groups studied the same subject matter, followed the same learning objectives, and received the same instructional duration.

Generative AI-Supported Instruction and Research Instruments

The intervention was implemented over eight meetings during a four-week period, with each meeting lasting approximately 90 minutes. The experimental group participated in learning supported by generative AI platforms, including ChatGPT and other AI-assisted educational tools. Generative AI was positioned as instructional scaffolding rather than as a provider of final answers. It provided temporary assistance through guiding questions, alternative explanations, examples, immediate feedback, information recommendations, and reflective prompts.

Students remained responsible for evaluating the accuracy and credibility of AI-generated information, constructing arguments, solving learning problems, and formulating conclusions. The teacher guided students in using AI responsibly and aided when students encountered misconceptions or difficulties in evaluating generated information.

Table 2. Stages of Generative AI-Supported Instruction

Learning stage	Student activity	Scaffolding function
Problem identification	Students identified learning problems and formulated relevant questions	Activating prior knowledge and directing inquiry
AI-assisted exploration	Students explored concepts and alternative explanations using generative AI	Providing information, examples, and initial guidance
Information verification	Students evaluated the accuracy, relevance, logic, and credibility of AI-generated information	Supporting analysis and evidence evaluation
Discussion and problem-solving	Students compared perspectives, constructed arguments, and solved problems collaboratively	Supporting reasoning and peer interaction
Reflection and revision	Students reflected on feedback and revised their explanations or conclusions	Supporting metacognitive reflection and improvement

These activities were designed to stimulate five dimensions of critical thinking: interpretation, analysis, evaluation, inference, and explanation. They were also intended to increase active participation, classroom discussion, analytical involvement, problem-solving engagement, and learning motivation. The control group received conventional instruction consisting of lectures,

textbook-based learning, teacher explanations, classroom questioning, and individual assignments without generative AI support.

Data were collected using a critical thinking test, a classroom engagement observation sheet, and documentation. The critical thinking instrument consisted of 20 essay and structured-response items developed according to interpretation, analysis, evaluation, inference, and explanation. The test was administered before and after the intervention, and students' responses were assessed using an analytical scoring rubric.

The instrument was validated by three specialists in educational assessment and evaluation. The content-validity results produced Aiken's V coefficients ranging from 0.82 to 0.94, while the reliability test produced a Cronbach's Alpha coefficient of 0.89, indicating high internal consistency.

Classroom engagement was assessed through structured observation focusing on active participation, classroom discussion, analytical activities, problem-solving activities, and learning motivation. These indicators represented behavioral, cognitive, and emotional engagement.

Table 3. Dimensions and Indicators of Classroom Engagement

Dimension	Observable indicators
Behavioral engagement	Active participation and involvement in classroom discussion
Cognitive engagement	Analytical activities and problem-solving involvement
Emotional engagement	Learning motivation, enthusiasm, and persistence

Each indicator was assessed using the same observation criteria in the experimental and control groups. The observation score was converted into a percentage using the following formula:

$$\text{Engagement Percentage} = \frac{\text{Observed Score}}{\text{Maximum Possible Score}} \times 100$$

Documentation included lesson plans, attendance records, students' assignments, records of AI-assisted activities, classroom photographs, students' written arguments, and teacher implementation notes.

Data Collection and Analysis

Data collection involved preparation, pretest administration, instructional intervention, classroom observation, and posttest administration. During the preparation stage, the researchers developed lesson plans, AI-supported instructional activities, assessment instruments, observation sheets, and scoring rubrics. Both groups completed the critical thinking pretest before the intervention. The experimental group then received generative AI-supported instruction, while the control group received conventional learning. Classroom engagement was observed throughout the eight meetings. After the intervention, both groups completed the critical thinking posttest.

The data were analyzed using descriptive and inferential statistics with SPSS version 26. Descriptive statistics included the mean, standard deviation, minimum score, maximum score, and percentage score. The Shapiro–Wilk test was used to examine data normality, while Levene's test was used to examine homogeneity of variance. A significance value greater than 0.05 indicated that the statistical assumption was satisfied.

Students' improvement in critical thinking was calculated using normalized gain analysis:

$$N\text{-Gain} = \frac{\text{Posttest Score} - \text{Pretest Score}}{\text{Maximum Score} - \text{Pretest Score}}$$

N-Gain scores above 0.70 were categorized as high, scores from 0.30 to 0.70 as moderate, and scores below 0.30 as low.

An independent-samples *t*-test was used to determine whether significant differences existed between the experimental and control groups. The significance level was set at 0.05. A significance value below 0.05 indicated a statistically significant difference. Cohen's *d* was calculated to determine the magnitude of the treatment effect, with values of approximately 0.20, 0.50, and 0.80 interpreted as small, medium, and large effects, respectively.

Classroom engagement was analyzed descriptively by comparing the percentage scores of the experimental and control groups. Behavioral engagement was calculated from active participation and classroom discussion, cognitive engagement from analytical and problem-solving activities, and emotional engagement from learning motivation. Because the engagement data were obtained from aggregate classroom observations rather than individual pretest–posttest measurements, they were interpreted as supporting descriptive evidence rather than as inferential causal findings.

The study hypothesized that students participating in generative AI-supported instruction would demonstrate significantly higher critical thinking skills than students participating in conventional instruction. It was also expected that the experimental group would demonstrate higher classroom engagement than the control group. The critical thinking hypothesis was tested inferentially, while classroom engagement was examined descriptively through observation data.

Ethical Considerations

Permission to conduct the research was obtained from the school principal and relevant educational authorities. Students and teachers were informed about the research objectives and procedures. Participation was voluntary, participants' identities were anonymized, and all data were treated confidentially. Generative AI was used under teacher supervision, and students were instructed to examine the accuracy, credibility, relevance, and ethical implications of AI-generated information

Results and Discussion

Implementation of Generative AI-Supported Instruction

Generative AI-supported learning was implemented in the experimental group through inquiry-based, interactive, and student-centered activities. Students used ChatGPT and other AI-assisted tools to explore information, analyze problems, verify AI-generated responses, construct arguments, participate in discussions, and revise conclusions based on feedback. The control group studied the same material through lectures, textbook-based activities, classroom questioning, and individual assignments without AI support.

Critical Thinking Outcomes

The descriptive results showed that both groups improved after the instructional intervention. However, the experimental group achieved a greater increase than the control group.

Table 4. Descriptive Statistics of Critical Thinking Skills

Group	N	Pretest Mean	Posttest Mean	Increase
Experimental group	32	85.40	93.75	8.35
Control group	32	72.15	78.30	6.15

The experimental group's mean score increased from 85.40 to 93.75, while the control group's mean score increased from 72.15 to 78.30. These results indicate that students who participated in generative AI-supported learning demonstrated stronger post-intervention critical thinking performance.

Normality and Homogeneity Tests

The Shapiro–Wilk normality test showed that all pretest and posttest significance values were greater than 0.05. The data were therefore normally distributed. Levene's test also produced significance values greater than 0.05, indicating homogeneous variances between the experimental and control groups.

Table 5. Results of the Assumption Tests

Test	Data	Experimental	Control	Conclusion
Shapiro–Wilk	Pretest	0.087	0.094	Normal
Shapiro–Wilk	Posttest	0.076	0.081	Normal
Levene's test	Pretest	0.118	0	Homogeneous
Levene's test	Posttest	0.103	0	Homogeneous

Because all values were greater than 0.05, the data met the assumptions required for parametric statistical analysis.

Improvement, Hypothesis Testing, and Effect Size

Normalized gain analysis was used to compare the improvement achieved by the two groups.

Table 6. N-Gain, Hypothesis Test, and Effect Size

Statistical indicator	Experimental group	Control group	Statistical value	Interpretation
Mean N-Gain	0.57	0.22	0	Experimental: moderate; Control: low
Posttest difference	93.75	78.30	($p=0.001$)	Significant difference
N-Gain difference	0.57	0.22	($p<0.001$)	Significant difference
Effect size	0	0	Cohen's ($d=0.87$)	Large effect

The experimental group obtained an N-Gain score of 0.57, categorized as moderate, whereas the control group obtained 0.22, categorized as low. The independent-samples t -test indicated significant differences in posttest scores and learning gains because both significance values were below 0.05. Cohen's $d = 0.87$ indicated that the instructional effect was large and educationally meaningful.

Critical Thinking Dimensions

Students' performance was also examined across interpretation, analysis, evaluation, inference, and explanation.

Table 7. Critical Thinking Skills by Indicator

Critical thinking indicator	Experimental group (%)	Control group (%)	Difference
Interpretation	91	79	12

Analysis	93	76	17
Evaluation	90	75	15
Inference	92	77	15
Explanation	94	80	14

The experimental group achieved higher scores across all critical thinking dimensions. The highest achievement was recorded for explanation at 94%, followed by analysis at 93% and inference at 92%. The largest difference between the groups occurred in analysis, indicating that AI-supported learning particularly strengthened students' ability to examine information, identify relationships, and evaluate alternative perspectives.

The manuscript contains another indicator table reporting slightly different experimental-group values. To maintain direct comparability between the experimental and control groups, Table 7 uses the comparative indicator data reported for both groups.

Classroom Engagement

Classroom observations showed that the experimental group demonstrated consistently higher engagement than the control group.

Table 8. Classroom Engagement During the Learning Process

Engagement indicator	Experimental group (%)	Control group (%)	Difference
Active participation	91.25	73.40	17.85
Classroom discussion	89.40	71.20	18.20
Analytical activities	93.10	68.85	24.25
Problem-solving activities	92.75	70.15	22.60
Learning motivation	94.05	75.40	18.65
Overall mean	92.11	71.80	20.31

The highest engagement score in the experimental group was learning motivation at 94.05%, followed by analytical activities at 93.10% and problem-solving activities at 92.75%. The greatest difference between the groups occurred in analytical activities, followed by problem-solving activities.

The indicators can also be grouped into three broader engagement dimensions.

Table 9. Classroom Engagement by Dimension

Dimension	Experimental group (%)	Control group (%)	Difference
Behavioral engagement	90.33	72.30	18.03
Cognitive engagement	92.93	69.50	23.43
Emotional engagement	94.05	75.40	18.65

Cognitive engagement demonstrated the largest difference. This finding suggests that generative AI-supported instruction was particularly effective in encouraging students to analyze information, investigate problems, and construct solutions. Because these data were derived from aggregate classroom observations, they are interpreted as descriptive supporting evidence rather than as inferential proof of causality.

Generative AI as Instructional Scaffolding for Critical Thinking

The findings demonstrate that generative AI-supported instruction was associated with substantially stronger critical thinking performance than conventional instruction. The experimental group achieved a posttest mean of 93.75, compared with 78.30 in the control

group. The experimental group also obtained a moderate N-Gain of 0.57, whereas the control group achieved a low N-Gain of 0.22. The difference in posttest scores was statistically significant at $p = 0.001$, while the difference in N-Gain scores was significant at $p < 0.001$. Furthermore, Cohen's $d = 0.87$ indicated a large instructional effect. These results suggest that the contribution of generative AI was not only statistically significant but also educationally meaningful.

The effectiveness of the intervention can be interpreted through the concept of instructional scaffolding. In this study, generative AI did not function merely as an answer-generating tool (Hwang & Yun-Fang Tu, 2021; Zawacki-Richter et al., 2019). Instead, it provided temporary cognitive support through guiding questions, alternative explanations, examples, immediate feedback, and access to multiple perspectives (Arseven & Bal, 2025; Putri et al., 2025). Such assistance enabled students to approach complex learning problems with greater structure while retaining responsibility for evaluating information and constructing final conclusions (Schmidt et al., 2025).

A central feature of the intervention was that students were not permitted to accept AI-generated responses uncritically. They were required to examine the accuracy, relevance, logic, and credibility of the information provided, compare it with other sources, discuss alternative interpretations, and revise their conclusions (Wiranatha et al., 2026). The learning sequence therefore positioned AI-generated content as material for inquiry rather than as unquestionable knowledge (Hikmat, 2022). The intervention included problem identification, AI-assisted exploration, information verification, collaborative argumentation, problem-solving, and reflective revision.

These activities encouraged students to move beyond memorization and routine task completion toward higher-order cognitive processing. Students had to determine what information was relevant, identify relationships among concepts, evaluate evidence, recognize inconsistencies, draw logical conclusions, and articulate reasoned explanations. Generative AI therefore supported the processes through which critical thinking was practiced without replacing students' intellectual effort (Arsyad et al., 2024; Kalyani, 2024).

This distinction is important because the educational value of generative AI depends on how it is pedagogically structured. Unrestricted use of AI to produce final answers may reduce cognitive engagement and encourage dependence. In contrast, structured use that requires verification, justification, and revision can transform AI into a scaffold for deeper reasoning (Bolarinwa, 2026). The present findings therefore suggest that generative AI becomes pedagogically valuable when it supports, rather than substitutes for, students' analytical and reflective processes.

Development of Critical Thinking Dimensions

The experimental group achieved higher scores than the control group across all five dimensions of critical thinking: interpretation, analysis, evaluation, inference, and explanation. The experimental group obtained scores of 91% for interpretation, 93% for analysis, 90% for evaluation, 92% for inference, and 94% for explanation. In comparison, the control group achieved 79%, 76%, 75%, 77%, and 80%, respectively.

Explanation was the highest-performing dimension in the experimental group. This result indicates that AI-supported learning strengthened students' ability to communicate ideas, justify arguments, and present conclusions logically (Arseven & Bal, 2025; Putra et al., 2025). During the intervention, students were repeatedly required to formulate responses, defend their reasoning, and revise explanations based on feedback from AI, teachers, and peers. These

activities likely helped students organize their thoughts more coherently and connect claims with supporting evidence (Bolarinwa, 2026; Sitepu et al., 2026).

The strong achievement in analysis can be attributed to students' exposure to multiple explanations and alternative perspectives. AI-assisted exploration enabled students to compare information, identify conceptual relationships, examine differences among responses, and determine which ideas were most relevant to the learning problem (Marsam et al., 2026; Tashtoush et al., 2025). The analytical process was therefore embedded within the interaction between students and AI-generated information.

Inference also developed strongly because students were required to synthesize information and draw conclusions from available evidence. Rather than merely repeating AI-generated statements, students had to decide which information could reasonably support a conclusion. This process encouraged them to distinguish between assumptions, evidence, and logical consequences (Budiman & Arif, 2018; Puling et al., 2024).

The improvement in evaluation is particularly significant in the context of generative AI. AI systems may produce responses that are inaccurate, incomplete, biased, or insufficiently contextualized (Aprianto et al., 2025; Chotim dkk, 2016; Sudarmono et al., 2025). Students were therefore explicitly instructed to assess the credibility, relevance, and logical consistency of AI-generated information. The instructional design transformed a potential limitation of generative AI into an opportunity for critical evaluation (Angel Sophia Intan et al., 2025; Susanto et al., 2025; Wahyuningsih et al., 2020). Instead of treating possible inaccuracies only as technological weaknesses, the learning activities used them as prompts for evidence checking, source comparison, and reflective judgment (Cantona et al., 2023; Straser et al., 2026).

Interpretation also improved because students were required to clarify learning problems, identify the meaning of information, and reformulate questions before consulting AI. Effective interaction with generative AI depends on the ability to understand the problem and construct appropriate prompts (Ardimiati et al., 2026; Sitepu et al., 2026). Consequently, students practiced interpretation both before and after receiving AI-generated responses.

Overall, the higher performance across all five dimensions indicates that generative AI-supported instruction did not influence only one isolated aspect of critical thinking (Gabelaia, 2025; Pratama et al., 2025). It created an integrated learning environment in which students interpreted problems, analyzed information, evaluated evidence, drew inferences, and communicated explanations through a continuous cycle of inquiry and revision (Fitrah et al., 2025; Linda et al., 2025; Wahyuningsih et al., 2020).

Generative AI and Classroom Engagement

The classroom observation findings showed that students in the experimental group were more actively involved in the learning process than students in the control group. The overall classroom engagement percentage reached 92.11% in the experimental group, compared with 71.80% in the control group. The experimental group also achieved higher percentages in active participation, classroom discussion, analytical activities, problem-solving activities, and learning motivation.

Behavioral engagement was reflected in students' active participation and involvement in classroom discussion. Students interacted with AI-supported platforms, formulated questions, responded to learning problems, exchanged ideas with peers, and contributed to collaborative activities. The experimental group achieved 91.25% in active participation and 89.40% in classroom discussion, compared with 73.40% and 71.20% in the control group. These

differences suggest that the instructional structure reduced passive participation and encouraged students to assume a more active role in the learning process.

Emotional engagement was reflected in learning motivation. The experimental group achieved 94.05%, while the control group obtained 75.40%. Generative AI may have increased students' interest because it provided immediate responses, interactive assistance, and opportunities to explore questions independently. Students could obtain feedback during the learning process and revise their responses without waiting until the end of the lesson. This responsiveness may have strengthened curiosity, enthusiasm, and persistence when completing challenging tasks.

Cognitive engagement demonstrated the largest difference between the groups. The experimental group achieved 93.10% in analytical activities and 92.75% in problem-solving activities, compared with 68.85% and 70.15% in the control group. When grouped into broader dimensions, cognitive engagement reached 92.93% in the experimental group and 69.50% in the control group, representing a difference of 23.43 percentage points.

This pattern is particularly important because cognitive engagement represents more than observable participation. It reflects students' willingness to invest sustained intellectual effort in examining evidence, identifying inconsistencies, comparing alternatives, solving problems, and constructing conclusions (Agbon, 2024; Chan & Hu, 2023). The strong cognitive engagement observed in the experimental group may therefore help explain their higher critical thinking performance (Kaigeng Li, Hong Wu, 2024; Ng et al., 2025).

The relationship between classroom engagement and critical thinking appears to be mutually reinforcing. Active participation and discussion provide students with opportunities to articulate reasoning, respond to alternative perspectives, and evaluate arguments (European Innovation Council and SMEs Executive Agency, n.d.; Irina Carnat, 2024). At the same time, intellectually challenging tasks make classroom participation more meaningful (Fang et al., 2025; Fitrah et al., 2025). The observation findings similarly showed that higher involvement in analytical discussion and problem-solving accompanied stronger critical thinking outcomes.

Nevertheless, classroom engagement was measured through aggregate observation percentages rather than individual pretest and posttest scores. Therefore, these findings should be interpreted as descriptive evidence accompanying the intervention, not as definitive inferential proof that generative AI alone caused the observed differences in engagement.

Conclusion

This study concludes that generative AI can function effectively as instructional scaffolding to strengthen secondary school students' critical thinking and classroom engagement when it is integrated into structured activities involving problem identification, information exploration, verification, discussion, problem-solving, and reflection. Students in the experimental group achieved higher critical thinking performance than those in the control group, as indicated by a posttest mean of 93.75 compared with 78.30, a moderate N-Gain of 0.57 compared with a low N-Gain of 0.22, significant differences in posttest and N-Gain scores, and a large effect size of Cohen's $d = 0.87$. The experimental group also demonstrated higher achievement across interpretation, analysis, evaluation, inference, and explanation, as well as stronger classroom engagement in active participation, discussion, analytical activities, problem-solving, and learning motivation. These findings answer the research objective by demonstrating that generative AI-supported instruction is more effective than conventional learning in promoting critical thinking and is accompanied by higher classroom engagement. Therefore, teachers are advised to use generative AI as guided instructional support rather than as a source of final answers, while schools should strengthen digital infrastructure, teacher

competence, AI literacy, academic-integrity policies, and responsible technology-use guidelines. Further studies should involve larger and more diverse samples, apply stronger controls for initial group differences, and measure classroom engagement using validated individual-level instruments.

Acknowledgment

The authors express their sincere gratitude to the principal, teachers, and eighth-grade students of SMP Negeri 1 Kota Bima for their permission, cooperation, and active participation throughout the research process. Appreciation is also extended to the Faculty of Teacher Training and Education, Nggusuwaru University, for the academic and institutional support provided during the preparation and implementation of this study. The authors further acknowledge the contributions of the instrument validators, research assistants, and colleagues who provided valuable suggestions and technical assistance in improving the quality of this research.

References

- Agbon, G. (2024). Who speaks through the machine? Generative AI as discourse and implications for management. *Critical Perspectives on Accounting*, 100, 102761. <https://doi.org/10.1016/j.cpa.2024.102761>
- Agol, R., Osias, N., & Comon, J. D. (2024). Use of Digital Technology for Learners' Cognitive Skills and Academic Performance in Social Studies. *European Modern Studies Journal*, 8(5), 322–349. [https://doi.org/10.59573/emsj.8\(5\).2024.29](https://doi.org/10.59573/emsj.8(5).2024.29)
- Angel Sophia Intan, Nidya Fitri, Yulia Wiji Astika, & Teta Wismar. (2025). Digital Project-Based Learning in English Language Education: Instrument for Achieving Sustainable Development Goals. *Jurnal Iqra' : Kajian Ilmu Pendidikan*, 10(1), 249–269. <https://doi.org/10.25217/ji.v10i1.5152>
- Aprianto, I., Sofyan, Rahmawati, S., & Sufyadi, S. (2025). Artificial intelligence and its effects on critical thinking and problem-solving abilities in higher education. *Indonesian Journal of Educational Development (IJED)*, 6(3), 704–719. <https://doi.org/10.59672/ijed.v6i3.5030>
- Ardimiati, N. C., Syahrial, S., & Syarif, A. (2026). Implementing a Deep Learning Pedagogical Approach to Improve Conceptual Understanding and Critical Thinking in Grade 5 IPAS Ecosystem Learning: A Classroom Action Research Study. *Journal of Innovation and Research in Primary Education*, 5(1), 1364–1371. <https://doi.org/10.56916/jirpe.v5i1.3027>
- Arseven, T., & Bal, M. (2025). Critical literacy in artificial intelligence assisted writing instruction: A systematic review. *Thinking Skills and Creativity*, 57, 101850. <https://doi.org/10.1016/j.tsc.2025.101850>
- Arsyad, M., Fitroh, I., & Arifin, M. S. (2024). Transforming 21st Century Education: Analysing the Implementation of Technology in Teaching and Learning. *Jurnal Ilmiah Edukatif*, 10(2), 332–341. <https://doi.org/10.37567/jie.v10i2.3423>
- Ayanwale, M. A., & Omeh, C. B. (2026). AI-supported problem-based learning for enhancing computational thinking skills in STEM education. *Computers in Human Behavior: Artificial Humans*, 7, 100263. <https://doi.org/10.1016/j.chbah.2026.100263>

- Bagdonaite, J., & Degiene, V. (2025). Artificial Intelligence in Primary Education: A Systematic Literature Review 2020–2025. *Informatics in Education*, 24(4), 697–736. <https://doi.org/10.15388/infedu.2025.24>
- Bolarinwa, T. (2026). The impact of artificial intelligence on students' academic development, critical thinking, cognitive skills, and learning outcomes. *International Journal of Applied Resilience and Sustainability*, 2(2), 1097–1129. <https://doi.org/10.70593/deepsci.0202045>
- Budiman, & Arif, I. (2018). Pengembangan Perangkat Pembelajaran Biologi Model Inkuiri Terbimbing Berkarakter Budaya Lokal “Nggahi Rawi Pahu” Untuk Melatihkan Kemampuan Berpikir Kritis Siswa SMA. *BIOEDUKASI: Jurnal Pendidikan Biologi*, 9(2), 109–123.
- Cantona, I. G. E., Suastra, I. W., & Ardana, I. M. (2023). HOTS Oriented Problem-Based Learning Model : Improving Critical Thinking Skills and Learning Outcomes of Fifth Grade Students in Science Learning. *Thinking Skills and Creativity Journal*, 6(1), 19–26.
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00411-8>
- Chotim dkk. (2016). Pengaruh Model Pembelajaran Berbasis Proyek terhadap Kemampuan Kerjasama Anak Usia 5-6 Tahun di TK Saiwa Dharma Banyuning Tahun Ajaran 2015/2016. *Bioscientist: Jurnal Ilmiah Biologi*, 11(2), 908. <https://doi.org/10.33394/bioscientist.v11i2.8405>
- European Innovation Council and SMEs Executive Agency. (n.d.). *Artificial intelligence and copyright: use of generative AI tools to develop new content*. 2024. https://intellectual-property-helpdesk.ec.europa.eu/news-events/news/artificial-intelligence-and-copyright-use-generative-ai-tools-develop-new-content-2024-07-16-0_en
- Fang, X., Ng, D. T. K., Tam, W. T., & Yuen, M. (2025). Design mobile computational thinking-integrated mathematics lessons based on the 5E instructional model for primary students. *Educational Technology Research and Development*, 0123456789. <https://doi.org/10.1007/s11423-025-10513-x>
- Fitrah, M., Sofroniou, A., Setiawan, C., Widiastuti, W., Yarmanetti, N., Jaya, M. P. S., Panuntun, J. G., Arfaton, A., Beteno, S., & Susianti, I. (2025). The Impact of Integrated Project-Based Learning and Flipped Classroom on Students' Computational Thinking Skills: Embedded Mixed Methods. *Education Sciences*, 15(4), 1–20. <https://doi.org/10.3390/educsci15040448>
- Gabelaia, I. (2025). The Role of Artificial Intelligence, Learning Analytics, and Sustainability for Future-Ready Universities. *Sustainability*, 18(10), 4884. <https://doi.org/10.3390/su18104884>
- Georgopoulou, M. S., Troussas, C., Krouska, A., & Sgouropoulou, C. (2025). Digital Literacy in Higher Education: Examining University Students' Competence in Online Information Practices. *Computers*, 14(12), 528. <https://doi.org/10.3390/computers14120528>
- Hikmat, H. (2022). The Readiness of Education in Indonesia in Facing The Society Era 5.0. *Jurnal Basicedu*, 6(2), 2953–2961. <https://doi.org/10.31004/basicedu.v6i2.2526>

- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in Education. Promise and Implications for Teaching and Learning*. Center for Curriculum Redesign.
- Hwang, G.-J., & Yun-Fang Tu. (2021). Roles and Research Trends of Artificial Intelligence in Mathematics Education: A Bibliometric Mapping Analysis and Systematic Review. *Journals Mathematics*, 9(6), 584. <https://doi.org/10.3390/math9060584>
- Ikram, M., Hanefar, S. B. M., Saleem, S. M. U., & Zulfiqar, F. (2026). Artificial intelligence in education: a systematic review of personalized learning trends and future directions. *Front. Educ*, 13(11), 1782626. <https://doi.org/10.3389/feduc.2026.1782626>
- Irina Carnat. (2024). Addressing the risks of generative AI for the judiciary: The accountability framework(s) under the EU AI Act. *Computer Law & Security Review*, 55, 106067. <https://doi.org/https://doi.org/10.1016/j.clsr.2024.106067>
- Kaigeng Li, Hong Wu, Y. D. (2024). Copyright protection during the training stage of generative AI: Industry-oriented U.S. law, rights-oriented EU law, and fair remuneration rights for generative AI training under the UN's international governance regime for AI. *Computer Law & Security Review*, 55, 106056. <https://doi.org/https://doi.org/10.1016/j.clsr.2024.106056>
- Kalyani, D. L. K. (2024). The Role of Technology in Education: Enhancing Learning Outcomes and 21st Century Skills. *International Journal of Scientific Research in Modern Science and Technology*, 3(4), 5–10. <https://doi.org/10.59828/ijrmst.v3i4.199>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., 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., ... 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>
- Linda, E., Rakhmawati, A., & Paidi. (2025). Development of Discovery-Based Outdoor Learning Integrated with Papua Local Wisdom to Improve Cooperation and Understanding of the Biodiversity Concept of Students of Phase E of Sman Khas Papua. *Jurnal Penelitian Pendidikan IPA*, 11(8), 328–336. <https://doi.org/10.29303/jppipa.v11i8.12140>
- Marsam, T., Afriani, R., Hidayah, R. N., Sipahutar, R. P., Uki, S., & Suhendri, H. (2026). The Integration of Deep Learning and Problem-Based Learning to Enhance Critical Thinking in Elementary Science Education. *Jurnal Penelitian Pendidikan IPA*, 12(2), 182–187. <https://doi.org/10.29303/jppipa.v12i2.13673>
- Moneus, Ahmed Mohammed, Y. S. (2024). Artificial intelligence and human translation: A contrastive study based on legal texts. *Heliyon*, 10(6), 1–14. <https://doi.org/https://doi.org/10.1016/j.heliyon.2024.e28106>
- Nacheva, R. (2024). Hired by Artificial Intelligence : Digital Inclusion Practices for People With Disabilities. *Journal of Underrepresented and Minority Progress*, 8(2), 10–41. <https://doi.org/10.32674/66c03968>
- Ng, D. T. K., Chan, E. K. C., & Lo, C. K. (2025). Opportunities, challenges and school strategies for integrating generative AI in education. *Computers and Education: Artificial Intelligence*, 8, 100373. <https://doi.org/10.1016/j.caeai.2025.100373>

- Pratama, A., Amelia, E., Nur'Aeni, F., Adhwa, L. N., Wibowo, N. A., Sabila, W. A., & Sundari, N. (2025). Menciptakan Pengelolaan Kelas Efektif Anak Usia Dini Dengan Gangguan Hiperaktif Dalam Pembelajaran di Taman Kanak-Kanak. *Kumarottama: Jurnal Pendidikan Anak Usia Dini*, 5(1), 59–73. <https://doi.org/10.53977/kumarottama.v5i1.2600>
- Puling, H., Manilang, E., & Lawalata, M. (2024). Logika dan Berpikir Kritis : Hubungan dan Dampak Dalam Pengambilan Keputusan. *Sinar Kasih: Jurnal Pendidikan Agama Dan Filsafat*, 2(2), 164–173.
- Putra, Z. H., Sumadinata, H., Witri, G., Barokah, R. G. S., Gunawan, Y., Dahnlisyah, & Mohd. Jamil, M. R. (2025). The effect of gender, high school major, admission process, and year of study towards pre-service elementary teachers' computational thinking skills. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-025-13429-4>
- Putri, M. W., Irwan, I., & Rahmadani, S. (2025). Pengaruh Literasi Digital Terhadap Kemampuan Berpikir Kritis dan Berpikir Kreatif Siswa pada Pembelajaran Sosiologi di SMAN 7 Solok Selatan. *Pendas : Jurnal Ilmiah Pendidikan Dasar*, 10(3), 221–231. <https://doi.org/10.23969/jp.v10i03.33081>
- Saar, M., Ljalikova, A., Theophilou, E., Alkhasawneh, S. N., & Hernández-Leo, D. (2025). Empowering Teachers to Enhance Student High Order Thinking Skills with Artificial Intelligence Educational Tools. *Two Decades of TEL. From Lessons Learnt to Challenges Ahead*, 442–456.
- Schmidt, D. A., Alboloushi, B., Thomas, A., & Magalhaes, R. (2025). Integrating artificial intelligence in higher education: perceptions, challenges, and strategies for academic innovation. *Computers and Education Open*, 9, 100274. <https://doi.org/10.1016/j.caeo.2025.100274>
- Sitepu, M. S., Prasajo, L. D., Hermanto, Salido, A., Nurhakim, L., Setyorini, E., Disnawati, H., & Wiratsongko, B. (2026). Mapping and Exploring Strategies to Enhance Critical Thinking in the Artificial Intelligence Era: A Bibliometric and Systematic Review. *European Journal of Educational Research*, 15(1), 305–322. <https://doi.org/10.12973/eu-jer.15.1.305>
- Straser, O., Bašić, M., Doorman, M., Weinberg, L., Kapelari, S., Katja Maaß 1Oliver Straser 1,* ,†ORCID, Matija Bašić 2,†, Michiel Doorman 3,†ORCID, L. W., Kapelari, S., & Maaß, K. (2026). Fostering Critical Thinking in STEM Education. *Education Sciences*, 16(3), 461. <https://doi.org/10.3390/educsci16030461>
- Sudarmono, M. A., Hasan, & Halima. (2025). Deep Learning Approach in Improving Critical Thinking Skills of Elementary School Students. *Jurnal Penelitian Pendidikan IPA*, 11(8), 60–70. <https://doi.org/10.29303/jppipa.v11i8.11708>
- Supriadi, N., Putra, F. G., Cahyaningtyas, E., & Imama, K. (2024). Optimizing Mathematical Representation Skills: Unveiling the Synergy between Quantum Teaching and Self-Esteem. *Journal of Philology and Educational Sciences*, 3(1), 34–42. <https://doi.org/10.53898/jpes2024313>
- Susanto, Kristiana, A. I., Dafik, Fatahillah, A., Venkatachalam, M., & Tanna, D. (2025). Enhancing Students' Conjecturing Skills Through RBL-STEM with Antimagic Coloring and Geometric Transformation in Batik Design. *Emerging Science Journal*, 9(2), 868–889. <https://doi.org/10.28991/ESJ-2025-09-02-019>

- Tan, X., Cheng, G., & Ling, M. H. (2025). Artificial intelligence in teaching and teacher professional development: A systematic review. *Computers and Education: Artificial Intelligence*, 8, 100355. <https://doi.org/10.1016/j.caeai.2024.100355>
- Tashtoush, M. A., Qasimi, A. B., Sheerawi, N. Al, & Hussein, L. A. (2025). The Efficacy of Utilizing Artificial Intelligence Techniques in Developing Critical Thinking in Mathematics among Secondary School Students and their Attitudes Toward it. *Iraqi Journal for Computer Science and Mathematics*, 6(1), 37–50. <https://doi.org/10.52866/2788-7421.1231>
- UNESCO. (2021). *The Recommendation: Ethics of Artificial Intelligence*. <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>
- Votto, A. M., Valecha, R., Najafirad, P., & Rao, H. R. (2021). Artificial Intelligence in Tactical Human Resource Management: A Systematic Literature Review. *International Journal of Information Management Data Insights*, 1(2), 100047. <https://doi.org/10.1016/j.ijime.2021.100047>
- Wahyuningsih, S., Nurjanah, N. E., Rasmani, U. E. E., Hafidah, R., Pudyaningtyas, A. ., & Syamsuddin, M. . (2020). STEAM Learning in Early Childhood Education: A Literature Review. *International Journal of Pedagogy and Teacher Education (IJPTE)*, 4(1), 33–44. <https://doi.org/10.20961/ijpte.v4i1.39855>
- Wardani, I. S., & Fiorintina, E. (2023). Building Critical Thinking Skills of 21st Century Students through Problem Based Learning Model. *Jurnal Pendidikan Indones*, 12(3), 461–470.
- Wiranatha, P. A., Mahayanti, N. W. S., & Adnyayanti, N. L. P. E. (2026). The Double- Edged Impact of Artificial Intelligence on Students' Critical Thinking Skills in Indonesia: A Systematic Literature Review. *Journal of Education and Teaching Learning*, 8(1), 74–91. <https://doi.org/10.51178/jetl.v8i1.3210>
- Zablotska, L., Chernii, L., Meleshchenko, V., Olendr, T., & Tsar, I. (2026). Developing Students' Critical Thinking through AI-Enhanced Foreign Language Instruction. *Arab World English Journal (AWEJ)*, 3, 1–15. <https://doi.org/10.24093/awej/AI3.9>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education -where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 1–27. <https://doi.org/10.1186/s41239-019-0171-0>