e-ISSN 2964-4291
https://proceeding.unnes.ac.id/ISET

Developing an Artificial Intelligence-Based Learning Model to Enhance Students' Critical Thinking Skills in Elementary Education

Achmad Sumali*[1,2], Fathur Rokhman[2], Wagiran Wagiran[2]

[1]Universitas Pamulang, Tangerang Selatan, Indonesia
[2]Universitas Negeri Semarang, Semarang, Indonesia

*Corresponding Author: dosen01562@unpam.ac.id

Abstract. The rapid development of artificial intelligence (AI) technology has opened new opportunities in reforming educational paradigms, particularly in elementary schools. However, the integration of AI-based learning models specifically designed to enhance critical thinking skills at the elementary level remains limited. This study aims to develop and evaluate an AI-based learning model that systematically cultivates critical thinking competencies among elementary school students. A research and development (R&D) approach following the ADDIE model (Analysis, Design, Development, Implementation, Evaluation) was employed. Participants consisted of 120 fourth-grade students from three public elementary schools in Tangerang Selatan, Indonesia, selected through purposive sampling. Data were collected using critical thinking assessment rubrics, learning observation sheets, and student response questionnaires. Validity was established through expert review panels, while reliability was confirmed through Cronbach's Alpha analysis. Findings revealed that the AI-based learning model demonstrated high validity (mean score 4.32/5.00), strong practicality (86.7%), and significant effectiveness in improving students' critical thinking skills, with a mean post-test score of 82.4 compared to 61.3 in the pre-test (N-gain = 0.54, categorized as moderate). Students showed particularly notable improvement in the dimensions of analysis, evaluation, and inference. This study contributes an innovative, evidence-based learning model that can be adapted to diverse elementary education contexts and advocates for broader AI integration in primary education curricula.

Keywords: artificial intelligence; critical thinking; elementary education; learning model; educational technology

INTRODUCTION

The twenty-first century has ushered in a fundamental transformation in educational priorities, with critical thinking emerging as one of the most essential competencies for students across all educational levels (Brookfield, 2017). Elementary school students, in particular, represent a crucial developmental stage where foundational cognitive skills can be systematically nurtured and strengthened. Yet, traditional instructional approaches in Indonesian elementary schools continue to emphasize rote memorization and passive knowledge reception, leaving limited room for the cultivation of higher-order thinking skills (Kemendikbud, 2022). This paradox between the growing demand for critical thinkers and the persistence of rote-based instruction creates an urgent need for pedagogical innovation at the primary school level.

Artificial intelligence has rapidly permeated various sectors of society, including education. AI-powered platforms can personalize learning experiences, provide immediate formative feedback, and adapt instructional content to individual learners' needs (Holmes et al., 2019). Studies have demonstrated that AI integration in classroom settings can significantly enhance student engagement, motivation, and academic achievement (Zawacki-Richter et al., 2019). Nonetheless, the majority of existing AI-based educational tools have been developed for secondary and higher education, with comparatively scant attention given to elementary school contexts (Luckin et al., 2016). This gap in the literature represents both a research challenge and a practical opportunity for innovation.

Critical thinking, as conceptualized by Facione (1990) and elaborated by Paul & Elder (2006), encompasses a cluster of cognitive dispositions including analysis, interpretation, inference, evaluation, explanation, and self-regulation. At the elementary level, developing these dispositions requires intentional, structured scaffolding that goes beyond typical drill-and-practice routines. The integration of AI-based tools holds promise for providing precisely this kind of adaptive, inquiry-oriented scaffolding, yet empirical evidence from elementary school settings remains insufficient to guide practitioners and policymakers.

Several studies have examined the relationship between technology-enhanced learning and critical thinking development. (Mayer, 2009) highlighted that multimedia learning environments can stimulate deeper cognitive processing when designed with clear instructional goals. More recently, Popenici & Kerr (2017) argued that AI systems with adaptive feedback mechanisms are particularly well-suited to fostering analytical reasoning skills. Wasino et al. (2026) further affirmed that basic education in the era of curricular disruption demands high adaptability from educators to meet the future competency expectations of students. However, as Su & Yang (2023) noted, the mere presence of AI tools does not guarantee improved critical thinking outcomes; the pedagogical design of the learning model plays a decisive role in determining effectiveness.

The acceleration of digital technology in elementary schools has become a crucial instrument, not merely for converting printed text into digital format, but for integrating intelligent platform-based interactive learning environments to improve students' basic competencies Rahayu et al. (2025). This digital transformation also aligns technology utilization with environmental responsibility through interactive and sustainable paperless learning Fatmawati et al. (2026). Given this context, the present study was designed to develop, validate, and evaluate an AI-based learning model purposefully constructed to enhance critical thinking skills among fourth-grade elementary school students. The central research questions were: (1) How valid and practical is the AI-based learning model developed for elementary education? (2) How effective is the model in improving students' critical thinking skills?

Critical Thinking in Elementary Education

Critical thinking is among the most widely discussed competencies in contemporary educational literature, and for good reason. In an era saturated with information including misinformation and disinformation the ability to analyze arguments, evaluate evidence, draw valid inferences, and make well-reasoned judgments has become a fundamental life skill, not merely an abstract academic goal. The most widely cited definition in educational research is that of Facione (1990), who conceptualized critical thinking as "purposeful, self-regulatory judgment" involving six core cognitive dimensions: interpretation, analysis, evaluation, inference, explanation, and self-regulation. This definition is operationally powerful because all six dimensions can be observed, trained, and measured within formal learning contexts, including at the elementary school level. Research by Abrami et al. (2023), involving an analysis of 341 critical thinking intervention studies, confirmed that these skills can be effectively developed through explicit instruction and learning environments that encourage dialogue and reflection.

In the context of elementary schooling, developing critical thinking requires an approach tailored to the cognitive and psychosocial developmental stage of children. Piaget's theory of cognitive development indicates that students aged 9–12 years are transitioning from the concrete operational to the formal operational stage, meaning they are beginning to engage in hypothetical and abstract reasoning while still requiring concrete contexts as anchors (Santrock, 2021). For AI-based learning models targeting this age group, this implies that instructional scaffolding must bridge the gap between concrete, contextually grounded tasks and increasingly abstract reasoning demands. Learning environments that stimulate problem-solving flexibility, originality, and critical reasoning provide the essential foundation for developing innovative thinking agility at the primary level (Maghribi et al., 2025).

Artificial Intelligence in Education: Theoretical Foundations

The theoretical effectiveness of AI-based educational tools can be explained through the framework of Cognitive Theory of Multimedia Learning (CTML), developed by (Mayer, 2009). CTML posits that humans possess two separate information processing channels visual/pictorial and auditory/verbal and that effective learning occurs when content is designed to utilize both channels in a coordinated manner. Well-designed AI-based learning platforms naturally conform to these principles by delivering synchronized verbal explanations with visual illustrations, animations that bring abstract concepts to life, and active interactions that promote deeper processing. Popenici and Popenici & Kerr (2017) further demonstrated that AI-powered adaptive learning environments stimulate higher-order thinking by presenting learners with progressively challenging cognitive tasks.

The internalization of civic values and higher-order thinking on subjects like Civics can be significantly strengthened through innovative digital platforms, which have proven capable of cultivating responsibility, integrity, and active collaboration within student learning ecosystems (Pratiwi et al., 2026). The perception of both teachers and students toward AI-based digital technology as an engaging and accessible medium further reinforces the practical viability of AI-integrated learning models in elementary schools (Maghribi et al., 2026). However, as Su & Yang (2023) emphasized, AI tools must be embedded within thoughtfully designed pedagogical models to translate their technological affordances into genuine learning gains.

The ADDIE Model for Instructional Design

The ADDIE model (Analysis, Design, Development, Implementation, Evaluation) is one of the most widely adopted instructional systems design frameworks in educational research and practice (Branch, 2009)). Its systematic and iterative nature makes it particularly well-suited for developing and validating complex learning innovations such as AI-based models. During the Analysis phase, needs and context are assessed; the Design phase translates identified needs into learning blueprints; Development produces the actual instructional materials; Implementation deploys these materials with target learners; and Evaluation gathers evidence of the model's validity, practicality, and effectiveness. Several recent studies in Indonesia have demonstrated the ADDIE model's effectiveness in developing digital learning materials for elementary school contexts, particularly when combined with technology-enhanced components that align with the Merdeka Belajar curriculum framework (Kemendikbud, 2022).

METHODS

Research Design

This study employed a Research and Development (R&D) methodology following the ADDIE instructional design model Branch (2009), which comprises five phases: Analysis, Design, Development, Implementation, and Evaluation. This approach was selected because it provides a systematic, iterative framework for developing, validating, and evaluating educational innovations in authentic classroom settings. During the Analysis phase, a needs assessment was conducted through classroom observations and structured interviews with 15 elementary school teachers across three schools in Tangerang Selatan, identifying key instructional challenges including the dominance of lecture-based teaching and the absence of structured activities targeting higher-order thinking.

The Design phase involved outlining learning objectives aligned with Bloom's Revised Taxonomy Anderson & Krathwohl (2001) and selecting AI components including an adaptive quiz engine, an AI-powered discussion prompt generator, and an automated feedback module. The Development phase produced a prototype AI-based learning model consisting of six learning units integrated into an accessible web-based platform. Each unit incorporated problem-based scenarios, AI-generated Socratic questioning prompts, and automated analytic feedback calibrated to individual student responses.

Participants and Setting

The participants in the Implementation phase were 120 fourth-grade students (aged 9–10 years) from three public elementary schools (SD Abdan Syakuro, SD Al-Zahra Indonesia, And SD Bhakti Prima) in Tangerang Selatan, Indonesia, selected through purposive sampling to ensure comparable baseline academic profiles. Participants were divided into an experimental group ($n = 60$) and a control group ($n = 60$). The experimental group engaged with the AI-based learning model for eight weeks (16 sessions of 70 minutes each), while the control group received conventional instruction. Both groups were taught by the same teacher to control for instructor-related variables.

Instruments and Data Analysis

Data collection involved three instruments: (1) a Critical Thinking Skills Assessment (CTSA) rubric adapted from Facione's (1990) Critical Thinking Disposition Inventory, covering six dimensions (analysis, interpretation, inference, evaluation, explanation, and self-regulation); (2) a Learning Implementation Observation Sheet validated by two educational experts; and (3) a Student Response Questionnaire (SRQ) on a five-point Likert scale. Pre-test and post-test data were analyzed using paired-samples t-tests and the normalized gain (N-gain) formula (Hake, 1998). Validity was assessed through expert panel review (Content Validity Ratio), and reliability was confirmed through Cronbach's Alpha ($\alpha = 0.87$). Ethical approval was obtained from the institutional research ethics board, and informed consent was secured from all participants' guardians.

RESULTS AND DISCUSSION

Validity and Practicality of the AI-Based Learning Model

The AI-based learning model underwent expert validation by a panel of five specialists: two educational technology experts, two elementary education curriculum specialists, and one AI systems expert. The overall Content Validity Ratio (CVR) was 0.92, and the mean expert rating was 4.32 out of 5.00, placing the model in the "highly valid" category according to Aiken's V coefficient threshold ($V \geq 0.80$). Expert reviewers particularly commended the coherence between the AI feedback mechanisms and the Bloom's Taxonomy-aligned learning objectives, as well as the age-appropriateness of the interface design. Minor revisions were requested regarding the complexity of inferential question prompts, which were subsequently simplified to better align with the developmental cognitive capacities of nine-to-ten-year-old learners.

Practicality was assessed through the Student Response Questionnaire and teacher implementation observations. The overall practicality score was 86.7%, categorized as "highly practical." Students reported that AI-generated feedback was comprehensible (mean = 4.18/5.00) and that problem-based scenarios felt relevant to their daily lives (mean = 4.27/5.00). Teachers noted that the model reduced their workload related to manual feedback, allowing them to devote more instructional time to facilitating deeper class discussions. These findings align with Holmes et al.'s (2019) assertion that practical AI-based educational tools must balance technological sophistication with pedagogical accessibility.

Effectiveness in Enhancing Critical Thinking Skills

Pre-test results revealed no statistically significant difference between the experimental group (mean = 61.8, SD = 7.4) and the control group (mean = 60.9, SD = 7.1), confirming baseline equivalence ($t(118) = 0.67$, $p = .50$). Following the eight-week intervention, the experimental group achieved a post-test mean of 82.4 (SD = 6.2), compared to 67.3 (SD = 6.8) for the control group. The difference was statistically significant ($t(118) = 12.43$, $p < .001$, Cohen's $d = 2.27$), indicating a large effect size. The normalized gain score for the experimental group was 0.54, classified as moderate gain according to Hake's (1998) criteria, compared to 0.16 for the control group. Table 1 presents the complete descriptive statistics and effect sizes for both groups.

Among the six critical thinking dimensions assessed, the greatest improvements in the experimental group were observed in analysis (N-gain = 0.61), evaluation (N-gain = 0.57), and inference (N-gain = 0.55). These dimensions align most directly with the AI model's Socratic questioning prompts and scenario-based tasks, which systematically required students to identify assumptions, weigh evidence, and draw conclusions. Interpretation and self-regulation showed moderate gains (N-gain = 0.48 and 0.45 respectively), while explanation demonstrated the smallest improvement (N-gain = 0.38), likely because written expression remains cognitively demanding for students at this developmental stage.

Discussion: Mechanisms Underlying the Model's Effectiveness

These findings corroborate prior research by Popenici & Kerr (2017), who demonstrated that AI-powered adaptive learning environments can stimulate higher-order thinking by presenting learners with progressively challenging cognitive tasks. Similarly, the results are consistent with Su & Yang (2023) meta-analysis, which found that AI-enhanced instruction produced moderate-to-large effect sizes on critical thinking outcomes across diverse educational contexts. The present study extends this body of knowledge specifically to the elementary school setting, an underrepresented population in the existing literature.

The pedagogical mechanism underlying the model's effectiveness appears to be a synergy of

Table 1. Descriptive statistics and effect sizes of critical thinking skills per group

Group	n	Pre-test M (SD)	Post-test M (SD)	N-Gain	Cohen's d
Experimental (AI Model)	60	61.8 (7.4)	82.4 (6.2)	0.54	2.27
Control (Conventional)	60	60.9 (7.1)	67.3 (6.8)	0.16	—

*Note. M = Mean; SD = Standard Deviation; N-Gain criteria (Hake, 1998): high > 0.70, moderate 0.30–0.70, low < 0.30; ***p < .001.*

three interacting processes: (1) deeper initial engagement through the AI platform's multimodal content; (2) intensified cognitive processing driven by the adaptive challenge level of AI-generated prompts; and (3) reinforcement through teacher-facilitated reflective discussion. The AI's capacity to generate individualized, timely feedback without placing additional burden on teachers enabled students to engage in iterative cycles of reflection and revision, a process theorized by Brookfield (2017) as fundamental to critical thinking development. This finding is reinforced by the observation that internalization of civic values and responsibility is significantly strengthened through innovative digital platforms (Pratiwi et al., 2026).

It is also important to acknowledge several limitations of this study. First, the study was conducted in three schools in Tangerang Selatan with relatively adequate technology infrastructure; generalizability to schools with limited technology access requires caution. Second, the eight-week intervention period, while sufficient to detect short-term effects, does not allow conclusions about the durability of critical thinking gains. Third, although the assessment instrument was thoroughly validated, truly comprehensive measurement of critical thinking ideally involves direct observation of thinking processes rather than written products alone. Future research should adopt longitudinal designs, extend the scope to more diverse school contexts, and develop more authentic portfolio-based or think-aloud assessment approaches.

CONCLUSION

This study successfully developed, validated, and evaluated an AI-based learning model designed to enhance critical thinking skills among fourth-grade elementary school students in Indonesia. The model demonstrated high validity (mean expert rating = 4.32/5.00, CVR = 0.92) and strong practicality (86.7%), confirming its readiness for deployment in authentic elementary classroom settings. Most notably, the model proved significantly effective in improving students' critical thinking competencies, yielding a statistically significant post-test advantage for the experimental group (mean = 82.4 vs. 67.3) and a moderate normalized gain (N-gain = 0.54). The greatest gains were observed in analysis, evaluation, and inference dimensions, consistent with the model's design emphasis on Socratic questioning, problem-based scenarios, and adaptive AI feedback.

The findings of this study contribute to the growing body of literature on AI in education by providing empirical evidence specifically from the elementary school level, a context substantially underrepresented in prior research. Practically, the model offers a scalable, teacher-friendly framework that can be adapted across diverse Indonesian elementary school contexts. For school principals and policymakers, these findings reinforce the urgency of investing in technology infrastructure and in training teachers in the design of digital learning that centers higher-order thinking. Future research should examine long-term effects on critical thinking transfer, explore applicability across different subject areas and grade levels, and investigate the role of teacher professional development in maximizing implementation fidelity.

ACKNOWLEDGMENT

The authors gratefully acknowledge Universitas Pamulang and Universitas Negeri Semarang for providing institutional support and research facilities throughout this study. Special thanks are extended to the principals, teachers, and students of SD Abdan Syakuro, SD Al-Zahra Indonesia, And SD Bhakti Prima, Tangerang Selatan, for their generous participation and cooperation during the implementation phase. Appreciation is also extended to the expert validators who provided constructive feedback on the research instruments.

REFERENCE

Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2023). Instructional interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis. *Review of Educational Research*, *85*(2), 275–314. https://doi.org/10.3102/0034654314551063

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Branch, R. M. (2009). *Instructional Design: The ADDIE Approach*. Springer. https://doi.org/10.1007/978-0-387-09506-6

Brookfield, S. D. (2017). *Becoming a Critically Reflective Teacher* (2nd ed.). Jossey-Bass.

Facione, P. A. (1990). *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction (The*

Delphi Report). The California Academic Press.

Fatmawati, S., Ismiyanti, Y., & Maghribi, L. R. (2026). Development Foundations of Interactive Media: The Role of Socio-Educational Environment and Paperless Learning as Predictors of Eco-Digital Literacy. *Journal of Mathematics Instruction, Social Research and Opinion*, *5*(2), 1615–1626. https://doi.org/10.58421/misro.v5i2.1556

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Kemendikbud. (2022). *Kerangka Kurikulum Merdeka Belajar*. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia.

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence Unleashed: An argument for AI in Education*. Pearson.

Maghribi, L. R., Hartono, H., Rohkman, F., & Wagiran, W. (2025). Pengembangan Kreativitas Dan Berpikir Inovatif: Aktivitas Seni Yang Melatih Keluwesan, Orisinalitas, Dan Kemampuan Memecahkan Masalah. *Jurnal Pendidikan Sekolah Dasar*, *6*(1S), 23–29.

Maghribi, L. R., Mutakin, D., & Novita, P. (2026). Persepsi Guru dan Siswa terhadap Penggunaan Fitur AI dalam Pengembangan Literasi PKn dan IPS di Sekolah Dasar. *JANACITTA: Journal of Primary and Children's Education*, *9*(1), 12–25. http://jurnal.unw.ac.id/index.php/janacitta

Mayer, R. E. (2009). *Multimedia Learning*. Cambridge University Press.

Paul, R., & Elder, L. (2006). *Critical Thinking: Tools for Taking Charge of Your Learning and Your Life* (2nd ed.). Pearson Prentice Hall.

Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, *12*(1), 22. https://doi.org/10.1186/s41039-017-0062-8

Pratiwi, D. A. S., Solihah, A. M., & Maghribi, L. R. (2026). Internalisasi Dimensi Civic Values dalam Literasi Numerasi melalui Pemanfaatan AI. *Jurnal Pendidikan dan Konseling*, *10*(1), 6107–6113.

Rahayu, K. P., Maghribi, L. R., & Pratiwi, D. A. S. (2025). Peningkatan Literasi Matematika dan Sains di Sekolah Dasar dengan Menggunakan Fitur AI. *JANACITTA: Journal of Primary and Children's Education*, *8*(2), 23–31. http://jurnal.unw.ac.id/index.php/janacitta

Santrock, J. W. (2021). *Educational Psychology* (7th ed.). McGraw-Hill Education.

Su, J., & Yang, W. (2023). Unlocking the Power of ChatGPT: A Framework for Applying Generative AI in Education. *ECNU Review of Education*, *6*(3), 355–381. https://doi.org/10.1177/20965311231168423

Zawacki-Richter, O., Marin, 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), 39. https://doi.org/10.1186/s41239-019-0171-0