Online Learning in Educational Research

Volume 5, Issue 1, 143-154.
e_ISSN: 2798-2351
https://www.journal.foundae.com/index.php/oler/index
DOI: 10.58524/oler.v5i1.720

Deep Learning Pedagogies Enhance AI Literacy in Elementary Students: A Five-Cycle Implementation Study

Stefanus Christian Relmasira*
Satya Wacana Christian University
INDONESIA

Jonan Phillip Donaldson
University of Alabama at Birmingham
USA

Article Info

Article history:

Received: April 28, 2025
Revised: June 15, 2025
Accepted: June 19, 2025

Keywords:

AI Literacy;
AMAIL Model
Artificial Intelligence (AI);
Deep Learning;
Elementary Educatio.

Abstract

The urgency to integrate Artificial Intelligence (AI) literacy into primary education in Indonesia is driven by the increasing presence of AI technologies in everyday life and the nation's strategic vision to prepare a future-ready workforce. However, current teaching practices remain largely behavioristic and content-driven, lacking the pedagogical depth needed to foster conceptual understanding and ethical engagement with AI. This study addresses this gap by investigating how deep-learning pedagogies—approaches that pursue deep learning as a goal through active, reflective, and collaborative experiences— can be used to improve AI literacy among fifth-grade students. Grounded in design-based research (DBR), the study implemented and refined the Associative Model of AI Literacy (AMAIL), a framework integrating cognitive constructivism, social constructivism, constructionism, and transformative learning theories. The intervention spanned five cycles in three public schools in Salatiga, involving 118 students. Learning outcomes were assessed using pre- and post-tests and student reflections, with analysis conducted through bootstrap methods and Exact McNemar's tests. Findings showed statistically significant improvements in students' ability to recognize AI, explain its logic, and reflect on its ethical implications ($p < .001$). The study demonstrates how deep-learning approaches, when applied iteratively and contextually, can foster not only technical understanding but also critical and ethical AI literacy in primary education. These findings can inform educators and government stakeholders in designing and implementing pedagogical strategies that support comprehensive AI literacy development at the primary level

To cite this article: Relmasira, S. C., & Donaldson, J. P. (2025). Deep learning pedagogies enhance AI literacy in elementary students: A five-cycle implementation study. *Online Learning in Educational Research, 5*(1), 143-154. https://doi.org/10.58524/oler.v5i1.720

INTRODUCTION

The rapid development of artificial intelligence (AI) technology has transformed how humans interact with digital technology. The integration of AI in daily life, from content recommendations on social media to virtual assistants, creates a new urgency in education to prepare younger generations for the AI era (Long & Magerko, 2020; Relmasira et al., 2023). In Indonesia, awareness of the importance of this educational transformation is reflected in the *Merdeka Belajar* (Freedom to Learn) policy as regulated in the Minister of Education, Culture, Research, and Technology of the Republic of Indonesia regulation number 12 of 2024 concerning curriculum in early childhood education, elementary education, and secondary education levels, and the deep learning initiative launched by the Ministry of Elementary and Secondary Education to develop student competencies comprehensively and holistically (Pusat Kurikulum dan Pembelajaran, 2025).

* **Corresponding author:**
Stefanus Christian Relmasira, Satya Wacana Christian University, INDONESIA. ✉ srelmasira@uksw.edu

To address these demands, educators must go beyond surface-level instruction and adopt pedagogies that pursue deep learning as a goal through active, reflective, and collaborative experiences. Deep learning emphasizes a comprehensive and meaningful understanding through the intrinsically motivated actualization of competence and interest, rather than merely the surface learning of technical mastery or memorization (Biggs, 1987; Nasir et al., 2021; Sawyer, 2022; Trigwell & Prosser, (Biggs, 1987; Nasir et al., 2021; Sawyer, 2022; Trigwell & and Prosser, 1996). The approach aligns with the need for developing AI literacy that encompasses technical, social, and ethical aspects (Heyder & Posegga, 2021). In the elementary school context, deep learning becomes highly relevant as it can help students build a solid understanding of AI from an early age, while developing critical awareness about the impact of this technology on society (Touretzky et al., 2019; Yim & Su, 2025).

Deep Learning pedagogies are essential for improving AI literacy at the elementary level due to several key reasons. First, introducing AI concepts at a young age helps students develop a foundational understanding of AI technologies, which is crucial as AI becomes more integrated into daily life (Moon et al., 2024; Ottenbreit-Leftwich et al., 2023; Yang, 2022). Teaching AI literacy at an early age also helps students understand the ethical implications of AI, including issues related to data justice and disinformation (Voulgari et al., 2021; Yim, 2024). In this case, students are involved in a critical thinking process where students are learning deeply by critically analyzing AI technologies and their impact on society, hence fostering a more informed and responsible future generation (Ojeda-Bazaran et al., 2021; Voulgari et al., 2021). Furthermore, curricula like Primary AI literacy, which use constructivist and meaningful learning as part of deep learning pedagogy, have shown that students can effectively grasp AI concepts such as machine learning and computer vision (Ottenbreit-Leftwich et al., 2023). In comparison with this study, the implementation of AI literacy was conducted across multiple iterations and provides statistical evidence that emphasizes ethics and reflective practice.

Although research on AI literacy at the elementary education level is developing (Chung et al., 2025; Liu & Zhong, 2024), there remains a gap in understanding how deep learning can be effectively implemented to improve AI literacy among elementary school students. Previous research has focused more on technical aspects of AI learning (Wu et al., 2024) or on developing general AI literacy frameworks (Almatrafi et al., 2024). To address this gap, this study adopts the Associative Model of AI Literacy (AMAIL) model, which integrates principles of cognitive constructivist learning theory, social constructivist learning theory, constructionist learning theory, and transformative learning theory (Relmasira et al., 2023, 2024). This model provides a comprehensive framework for developing AI literacy that includes abilities in interacting with AI, recognizing AI, explaining and evaluating AI, and understanding the ethics of AI use.

The significance of this research becomes increasingly relevant considering the rapid development of generative AI, as demonstrated by Epstein et al. (2023) Gong et al. (2023). The ability of AI to generate content increasingly similar to human work creates a new urgency in education to prepare students for the challenges and opportunities of AI technology. Deep learning, with its emphasis on constructive and critical learning (diSessa, 2022; Kolodner et al., 2003), becomes a potential approach to develop comprehensive and meaningful AI literacy. This research aims to test the effectiveness of implementing deep learning to improve AI literacy capabilities among elementary school students.

METHOD

This research is part of a broader design-based research (DBR) study (Anderson & Shattuck, 2012) focused on AI literacy development for elementary school students. This article specifically presents the quantitative analysis results of the effectiveness of deep learning approaches in enhancing students' AI literacy. DBR was chosen because the iterative cycles design enables rapid refinement of the deep-learning activities in authentic Grade 5 classrooms, preserving ecological validity while allowing researchers and practitioners to co-analyze emergent evidence and adjust tasks, scaffolds, and assessments in real time. This approach is especially apt for Indonesia's newly devolved *Merdeka Belajar* curriculum, where adaptive, context-responsive learning designs are encouraged, and it aligns philosophically with deep-learning pedagogy itself, which values continuous inquiry, reflection, and knowledge building. To visualize the structure and stages of this iterative process, Figure 1 illustrates the design-based research framework employed in this study, detailing each phase from initial analysis to model refinement and classroom implementation.

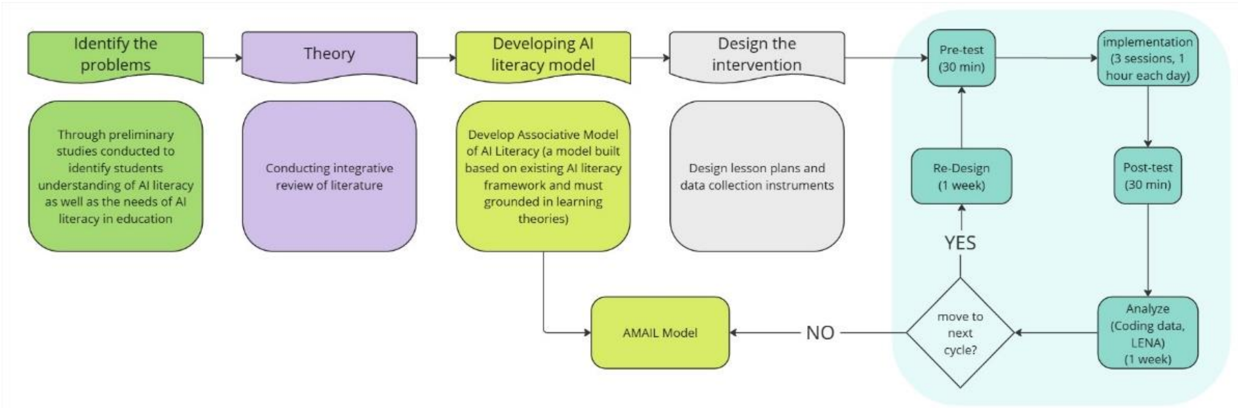


Figure 1. Design-Based Research Study for AI Literacy Development

The design begins with preliminary studies that surface pupils' AI-literacy gaps and teachers' needs, feeds these findings into an integrative literature review to ground a new Associative Model of AI Literacy (AMAIL). How this model was constructed has been explained in the previous publication (Relmasira et al., 2024). The model consists of 4 constructs and 16 competencies of AMAIL as described in Table 1.

Table 1. The 4 Constructs and the 16 Competencies of AMAIL

Construct	Competencies	Description
Recognition of AI	Recognize AI: Recognize AI (and not AI).	The ability to recognize when a digital technology is an AI, and when it is non-AI software.
	Recognize Products: Recognize products produced by AI (and not AI).	The ability to determine if a product was created by an AI or not.
Explaining and evaluating AI	Explain Limitations: Explain the limitations of AI.	The ability to explain what AI cannot do.
	Explain Strengths: Explain the strengths of AI.	The ability to explain areas in which AI is superior to non-AI technologies or human capabilities.
	Explain Roles: Explain the relationships between AI and humans.	The ability to explain the relationships between AI and humans, including how each interacts with the other and how each is impacted by the other.
	Explain Potential: Explain the potentials of AI.	The ability to describe what AI is able to do, including future potentials.
	Explain Data: Explain the data (or sensory inputs) that an AI uses.	The ability to explain the types of data (including sensory inputs) that AI uses.
	Explain Logic: Explain the logic (or procedures) that an AI uses.	The ability to explain in simple terms the way AI operates.
Interacting with AI	Solve Problems: Use AI to solve problems.	The ability to solve a variety of problems using AI, particularly with AI as a collaborator.
	Generativity: Use AI generatively (make stuff).	The ability to make new things using AI.
	Collaboration: Collaborate with others using AI.	The ability to leverage AI to more effectively collaborate with others.

Construct	Competencies	Description
	Art and Expression: Use AI for art and expression.	The ability to use AI to express yourself, including creation of art.
	Fun: Use AI for fun (entertainment).	The ability to have fun and be entertained through the use of AI.
Ethics Regarding AI	Evaluate Power: Evaluate who has power with any specific AI.	The ability to evaluate a specific AI in terms of who has power.
	Evaluate Impact: Evaluate who profits, benefits and harmed by any specific AI.	The ability to evaluate a specific AI in terms of who profits, who benefits, and who is harmed.
	Identify Purpose: Identify the purpose of any specific AI.	The ability to identify the explicit and implicit purposes of an AI.

We used AMAIL to craft lesson plans and collect data. In the intervention, each classroom cycle then runs a 30-minute pre-test, three one-hour lessons, and a 30-minute post-test, followed by a week-long analysis of work artefacts and learning experience network analysis (Donaldson et al., 2024) (LENA) of captured discourse that informs a rapid re-design before the next cycle.

Unlike classroom action research, which typically focuses on solving immediate instructional problems in a specific classroom context, DBR enabled us to systematically test, refine, and establish the effectiveness of the AMAIL model across multiple settings (Barab & Squire, 2004). This research required collaboration among diverse stakeholders, educational researchers, AI specialists, curriculum developers, and classroom teachers across three different schools to generate broadly applicable knowledge about deep learning approaches to AI literacy. The DBR methodology supported our need for robust quantitative measurement across five implementation cycles, allowing for progressive refinement of the intervention based on empirical evidence rather than solely practitioner reflection. This methodological choice was essential for developing generalizable principles that can inform AI literacy instruction across various elementary education contexts in Indonesia and beyond.

The effectiveness measurement employed a single-group pre-test post-test embedded within five cycles with different student groups in each cycle. Each cycle encompassed three learning sessions implementing the AMAIL model, focusing on four main constructs: interaction with AI, AI recognition, AI explanation and evaluation, and AI ethics. Rather than comparing AMAIL to an alternative method, the focus was on assessing the learning gains attributable to the intervention itself. This design is congruent with the overarching Design-Based Research (DBR) framework employed (Anderson & Shattuck, 2012; Barab & Squire, 2004). The pre- and post-test data collected in each of the five iterative cycles provided essential empirical feedback for evaluating the impact of the intervention at that stage and guiding pedagogical refinements for subsequent cycles. This approach offered a feasible and contextually appropriate method for investigating intervention effectiveness in the novel area of elementary AI literacy within authentic classroom settings, where establishing matched control groups can pose significant logistical and ethical challenges.

The research was conducted in three public elementary schools in Salatiga, Central Java, with a total of 118 fifth-grade students divided across 5 cycles. Sample selection was performed purposively, considering the availability of computer facilities and internet access at schools. Participant characteristics included students aged 10-11 years with varying access to technology at home. One of the instruments used in this study was a pre-test and post-test questionnaire with 6 Likert scale questions (1-5) to measure AI recognition, and 14 dichotomous questions to measure AI interaction, understanding, and ethics. The instrument's validity and reliability were tested with face validity and content validity by two experts. Reliability testing with SPSS showed Cronbach's alpha reliability (Taber, 2018) of p=0.731 for Likert scale questions, while the Cronbach's alpha reliability for dichotomous questions showed p=0.707.

Data collection was conducted through several stages:

1. Pre-test administration
2. Implementation of deep learning with the AMAIL approach in three sessions:
   - Session 1 - Basic interaction with AI
   - Session 2 - AI recognition and evaluation
   - Session 3 - AI ethics and impacts
3. Post-test administration
4. Collection of student reflection data

Data analysis was conducted using two approaches:

1. Bootstrap Analysis (Davison & Hinkley, 1997). This technique was used to analyze Likert scale data (questions 1-6). This approach was chosen because the data was not normally distributed. Significance level $\alpha = 0.05$
2. Exact McNemar's Test (McNemar, 1947). This technique was used to analyze dichotomous data (questions 7-20) and test response changes before and after intervention. The analysis used R software. Significance level $\alpha = 0.05$

The intervention of the deep learning approach implementation was conducted through three integrated sessions:

Session 1 - Basic Interaction with AI. In this session, students engaged in active, deep, and enjoyable learning. The session contained unplugged image categorization activities (to understand how AI performs data classification in Machine Learning), then students collaborated on projects using Google's Teachable Machine (Google, 2024a) for AI data training. The allocated time was approximately 60-90 minutes for session 1.

Session 2 - AI Recognition and Evaluation. In session 2, students conducted creative and enjoyable exploration activities. They learned to construct knowledge about how AI operates by creating creative works with Auto Draw (Google, 2024b). However, before that, they played a "guess the drawing" game with the AI Quick Draw application (Google, 2024c) to understand how AI predicts given input data. Indirectly, in session 2, students built an understanding of AI logic and its limitations. The allocated time was 60-90 minutes.

Session 3 - AI Ethics and Impacts. In session 3, students tried to collaborate in building awareness of AI ethics and its impacts on society. They engaged in discussion activities analyzing "deepfake" AI products that they could find on social media, then they discussed the social impacts of these AI technologies. For session 3, the allocated time was 60-90 minutes.

RESULTS AND DISCUSSION

The result and analysis, framed within the AMAIL framework, evaluated students' competencies in recognizing, interacting with, understanding, evaluating, and considering ethics in AI. The qualitative analysis of the result was described in an earlier publication (Relmasira et al., 2023), whereas the quantitative results from the Likert-scale responses indicated mixed outcomes. Most notably, for the question assessing students' perception of YouTube's personalized suggestions, there was a significant increase (mean difference 0.245, $p = .020$), demonstrating heightened awareness of AI personalization post-intervention. Conversely, other questions showed minimal or statistically insignificant changes, such as perceptions about decision-making by apps or computers (mean difference 0.106, $p = .266$) and AI intentions (mean difference 0.064, $p = .591$). For binary response questions analyzed using the Exact McNemar's test, significant improvements were consistently observed across several iterations. Specifically, in recognizing AI-generated images, substantial advancements were seen, with the final iterations yielding significant results (e.g., Q8: odds ratio = 6, $p = .0013$; Q9: odds ratio = 5, $p = .0386$). Questions regarding interactions with AI, such as using AI for entertainment or problem-solving, showed similarly significant improvements (e.g., Q11: iteration 4 odds ratio infinite, $p < .001$). Additionally, the intervention proved highly effective in enhancing students' ability to explain and evaluate AI, particularly regarding YouTube's recommendation logic (Q13: iteration 5 odds ratio infinite, $p < .000002$). In terms of AI ethics, notable advancements occurred in students' understanding of negative societal impacts of AI (Q20: iteration 4 odds ratio infinite, $p < .001$).

Collectively, these findings underscore the effectiveness of the AI literacy intervention, particularly in iterative stages, highlighting its progressive impact on students' comprehensive understanding and ethical considerations of AI technologies. Moreover, the iterative Learning-Experience Network Analysis revealed that the explaining & evaluating cluster mediated progress between hands-on interaction and ethical reflection, echoing but also elaborating on Zhang et al. (2022) the claim that critical explanation is a necessary bridge to ethical reasoning. Together, these data show that AMAIL fostered an integrated, theory-aligned trajectory from playful interaction to critical, ethical AI literacy, advancing the field by providing replicable evidence of how primary students can achieve deep, multi-dimensional AI understanding within regular school settings.

Table 2 presents the results of the Bootstrap Analysis measuring changes in students' understanding of various aspects of AI following the educational intervention. This analysis examines six key dimensions of AI understanding using a 5-point Likert scale, with mean differences indicating shifts in students' perceptions between pre-test and post-test assessments. Bootstrap methodology was employed due to the non-normal distribution of the data, providing a robust statistical approach for analyzing the intervention's impact. The analysis reveals varied outcomes across different aspects of AI understanding, with some dimensions showing positive changes and others demonstrating negative shifts. Particularly noteworthy is the significant improvement in students' understanding of AI personalization mechanisms (p=0.020), contrasted with the trend toward more realistic perceptions of application intelligence. These findings offer valuable insights into how the intervention influenced different facets of students' AI literacy, highlighting areas of successful knowledge construction as well as important conceptual recalibrations.

Table 2. Bootstrap Analysis Results for Understanding of AI (n=118)

Aspects of AI Understanding	Mean Difference	p-value	Cohen's d
AI Decision Making	0.106	0.266	-0.07
AI Intention	0.064	0.591	-0.07
Application Intelligence	-0.191	0.063	-0.18
Recommendation Fairness	0.085	0.368	-0.09
Recommendation Usefulness	-0.085	0.309	0.03
Recommendation Personalization	0.245	0.020*	0.19

Note: *p < 0.05

As shown in Table 2, the most notable improvement occurred in students' understanding of recommendation personalization, with a mean increase of 0.245 and a statistically significant result ($p = 0.020$). This indicates that students became more aware of how AI tailors content based on user preferences. Conversely, the mean difference for application intelligence was negative (-0.191, $p = 0.063$), suggesting a conceptual shift toward a more realistic understanding of AI capabilities. Similar modest or non-significant changes appeared across other constructs, such as AI decision making (0.106, $p = 0.266$) and recommendation usefulness (-0.085, $p = 0.309$). While the observed learning gains were small in magnitude, they reflect early stages of conceptual recalibration where students refine inflated or inaccurate beliefs about AI. This pattern aligns with the developmental nature of deep learning and the goals of AI literacy education. For reference, effect sizes calculated using Cohen's d were all below 0.2, indicating very small but meaningful shifts.

To complement the statistical significance analysis, Cohen's d values were calculated to estimate the magnitude of observed changes in students' AI understanding across all five implementation cycles. These effect sizes were consistently small ($|d| < 0.2$), with the largest effect observed in students' improved understanding of recommendation personalization ($d = 0.19$, $p = .020$). Because the intervention was progressively refined in each cycle and involved different student groups, these aggregated effect sizes should be interpreted as descriptive indicators of general learning trends rather than fixed treatment effects. In the context of design-based research, learning is understood as developmental and iterative rather than binary. Small shifts are expected and valued as signs of progressive refinement, especially in early phases of pedagogical innovation. Outcomes are interpreted within the evolving context of instructional design and conceptual change, not solely through statistical thresholds. From this perspective, the small effect sizes, particularly when aligned with gains in personalization understanding and qualitative evidence of ethical awareness, are seen as early indicators of conceptual recalibration toward more critical and realistic understandings of AI.

Figure 2 presents a visual representation of the bootstrap analysis results examining changes in students' understanding of AI after the deep learning intervention. The horizontal bars depict mean differences for each of the six AI understanding dimensions measured in the study, with positive values (extending right from the zero line) indicating improved understanding and negative values (extending left) suggesting a shift toward more critical or realistic perceptions. Statistical significance is highlighted by the green bar, while the color coding differentiates between positive (blue) and negative (red) changes in understanding.

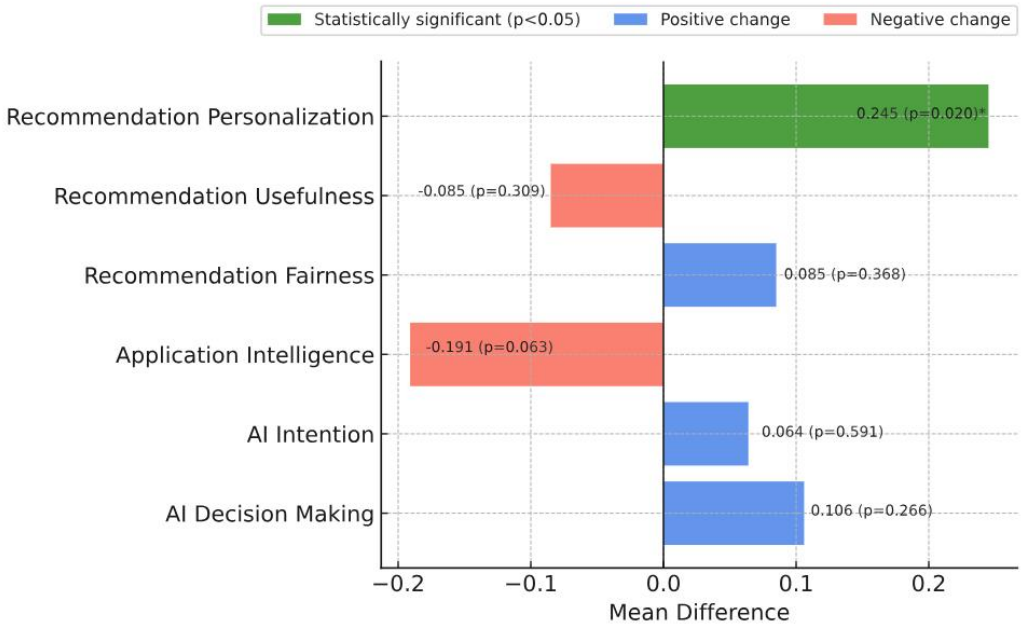


Figure 2. Bootstrap Analysis Results for Understanding AI

Figure 2 illustrates the magnitude and direction of changes in students' AI understanding, highlighting the significant improvement in Recommendation Personalization understanding (p=0.020) alongside the non-significant trends in other dimensions. This visualization reveals distinct developmental patterns across AI literacy components. The chart also highlights an important pattern not immediately apparent in the tabular data: the contrasting directionality of changes, with the three recommendation-related aspects showing distinctly different patterns (personalization improving significantly, fairness showing moderate improvement, and usefulness decreasing). This visualization reveals how students' conceptual understanding developed unevenly across different AI aspects, suggesting that the deep learning approach may have led to more nuanced and differentiated perceptions of AI systems rather than uniform improvements across all dimensions. This differentiated development likely occurred because the deep learning sessions explicitly addressed how AI personalizes content based on user data (Session 1 activities with Teachable Machine demonstrated data training), while simultaneously encouraging critical reflection on AI capabilities (Session 3 discussions on deepfakes and AI ethics). The decrease in perceived Application Intelligence and Recommendation Usefulness reflects a shift from potentially inflated initial perceptions toward a more realistic understanding of AI's limitations, representing a valuable educational outcome despite being statistically non-significant changes. The development of students' AI literacy in each learning iteration was analyzed using Exact McNemar's Test, with results summarized in Table 3.

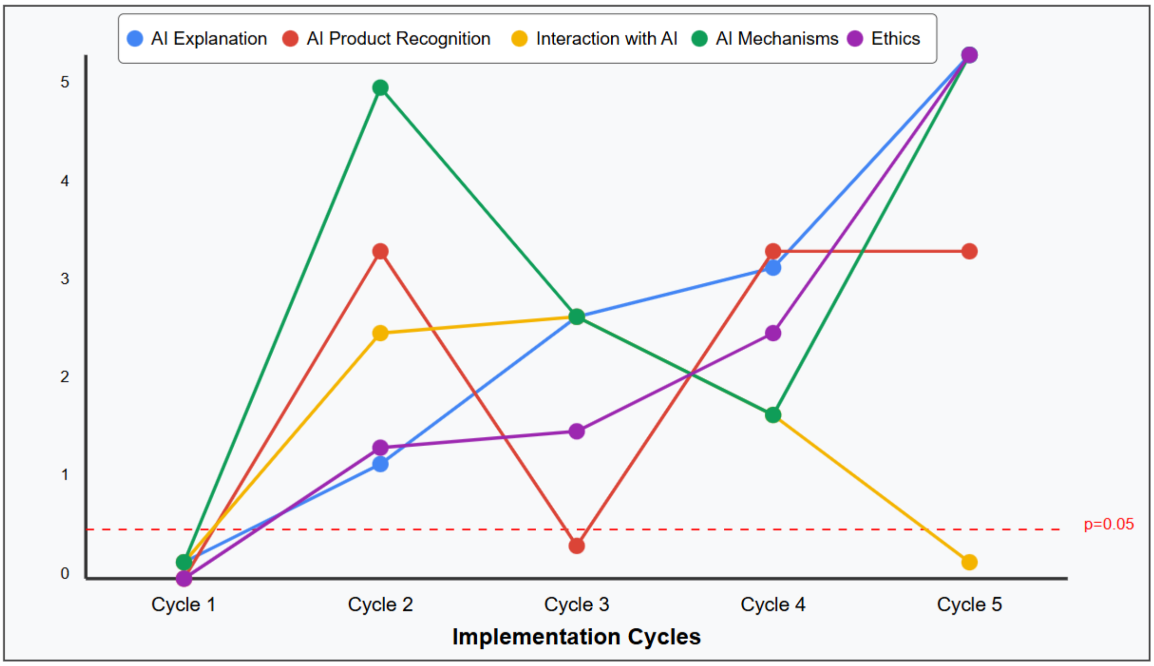
Table 3. Exact McNemar's Test Results for AI Literacy in Each Cycle

AI Literacy Aspect	1st Cycle	2nd Cycle	3rd Cycle	4th Cycle	5th Cycle
AI Explanation	p=0.109	p=0.01*	p=0.0002*	p=0.0009*	p<0.001*
AI Product Recognition	P=0.625	p=0.001*	p=0.179	p=0.001*	p=0.001*
Interaction with AI	P=0.125	p=0.003*	p=0.007*	p=0.015*	p=0.125
Understanding AI Mechanisms	P=0.125	p=0.000*	p=0.007*	p=0.015*	p<0.001*
AI Ethics Awareness	P=0.625	p=0.021*	p=0.031*	p=0.006*	p<0.001*

*Note: *p < 0.05*

The analysis results in Table 3 show the developmental progression in the effectiveness of deep learning implementation for enhancing AI literacy. In the first cycle, which was the implementation model trial stage, all AI literacy aspects showed non-significant results ($p>0.05$). This is understandable considering that the first cycle was a learning stage for instructors in implementing the deep learning model for AI literacy. Improvements in the learning model implementation are evident from the second cycle results, where all AI literacy aspects showed significant improvements ($p<0.05$). Students' abilities in explaining AI ($p=0.001$), recognizing AI products ($p=0.001$), interacting with AI ($p=0.003$), understanding AI mechanisms ($p=0.000$), and ethical awareness ($p=0.021$) showed significant improvements. This indicates that the adjustments and improvements from the first cycle reflection successfully enhanced learning effectiveness.

From the third to fifth cycles, the learning model implementation showed consistently significant results for most aspects, with some variations. For example, the ability to recognize AI products was not significant in the third cycle ($p=0.179$) but became significant again in the fourth and fifth cycles ($p=0.001$). These variations may be related to the characteristics of different student groups in each cycle. The highly significant improvements in the fifth cycle, particularly in AI explanation and ethical awareness aspects ($p<0.001$), indicate that the learning model had reached an optimal implementation level. This reflects the continuous refinement process in applying the deep learning model for AI literacy. This progression is further illustrated in Figure 3, which visualizes the changes in significance levels across the five implementation cycles for each aspect of AI literacy.



Note: Higher values indicate stronger statistical significance (p-values transformed as -log10(p))

Figure 3. Progression of Significance Levels Across Implementation Cycles

This pattern shows that the effectiveness of deep learning for AI literacy depends not only on the learning model itself but also on the maturity of its implementation. The consistent significant improvements in cycles after the trial stage indicate the success of the learning model refinement process through continuous reflection and improvement.

The research findings reveal interesting patterns in the effectiveness of deep learning implementation for enhancing elementary students' AI literacy. The quantitative analysis uncovers significant developmental progression across cycles, reflecting the refinement process in the learning model implementation.

In terms of basic AI understanding, a constructive shift in understanding occurred. Students developed a more realistic view of AI capabilities, shown by the negative change in perception of application "intelligence" ($p=0.063$). This indicates that deep learning helps students understand AI not as "magical" technology, but as a system with specific capabilities and limitations. This finding aligns with Chung et al. (2025) research emphasizing the importance of deconstructing AI misconceptions among elementary students. The significant improvement in understanding AI personalization ($p=0.020$) demonstrates the success of deep learning in developing students' contextual understanding. Students not only understood that AI can provide recommendations but also recognized that these recommendations are tailored to individual preference data. This understanding is essential as a foundation for developing critical AI literacy, as suggested by Long & Magerko (2020).

The developmental pattern from the first to the fifth cycle shows the process of refining the learning model implementation. The non-significant results in the first cycle do not indicate model failure, but rather reflect the adjustment process in learning implementation. The consistent significant improvements in subsequent cycles confirm the success of adjustments and improvements in applying the learning model. The ethical awareness aspect in AI literacy showed increasingly stronger improvements in later cycles, with the highest significance value in the fifth cycle ($p<0.001$). This indicates that as the learning model implementation matures, the ethical awareness aspect can be more effectively integrated with technical understanding. This finding strengthens Heyder & Posegga (2021) the argument about the importance of developing comprehensive AI literacy.

The varied progression of significance levels across implementation cycles, particularly the non-linear improvements observed in aspects like AI Product Recognition and Interaction with AI, reflects the complex nature of iterative educational design. These fluctuations can be attributed to several factors in the learning environment. Each cycle involved different student groups with varying prior experiences and learning preferences, naturally influencing their engagement with the intervention. Additionally, refinements made between cycles sometimes introduced new complexities while resolving previous issues, temporarily affecting certain aspects of AI literacy development. For instance, the decline in significance for AI Product Recognition in Cycle 3 ($p=0.179$) followed by improvement in Cycle 4 ($p=0.001$) demonstrates how specific instructional changes might have initially complicated this particular aspect before subsequent adjustments led to enhanced understanding. This pattern suggests that implementing effective AI literacy education is not a straightforward process but requires continued refinement responsive to students' evolving needs and challenges encountered during implementation.

The success of deep learning in enhancing AI literacy is inseparable from the approach that integrates hands-on experience, critical reflection, and contextualization in students' daily lives. This aligns with Papert's constructionist principles, emphasizing learning through creative projects and active exploration (Papert & Harel, 1991). The use of tools like Teachable Machine and unplugged activities provides concrete experiences that help students build a deeper understanding of AI. The results inform AI-literacy development in Indonesian elementary schools. A deep-learning pedagogy that combines iterative hands-on activities, guided explanation, and ethical reflection can be implemented within the national curriculum when technical skills and moral considerations are presented concurrently. These results are in line with the recommendations of Ma et al. (2025). These findings support the recommendation that teachers provide regular direct interaction with AI tools so students can relate abstract concepts such as training data, algorithmic bias, and personalized recommendations to concrete examples appropriate to their cognitive stage. These activities must be paired with discussions that require students to analyze both the capabilities and the limitations of AI, including fairness, privacy, and social impact. This integrated and developmentally appropriate design offers an empirically supported approach for promoting comprehensive AI literacy in primary education. While this research shows promising results, several limitations should be noted, including the relatively small sample size and context limited to elementary schools in one region. Further research on a larger scale and in more diverse contexts is needed to validate these findings.

LIMITATIONS

Despite these contributions, this study has several limitations. The relatively small and region-specific sample, along with the short duration of each implementation cycle, may have limited the depth and breadth of conceptual change achieved. While statistically significant gains were observed, the modest scale of these learning shifts suggests that longer-term interventions may be needed to foster deeper and more sustained AI literacy development. Addressing these constraints in future studies will be essential for expanding the impact and generalizability of AI literacy efforts.

CONCLUSION

This study shows that the Associative Model of AI Literacy (AMAIL) implemented through a deep-learning pedagogical approach defined by iterative cycles of hands-on exploration, guided explanation, and structured ethical reflection produced significant gains in Indonesian fifth-graders' ability to recognize AI outputs, explain core concepts such as personalization, and discuss social-ethical implications, with improvement in understanding AI achieved over five implementation cycles. This indicates that the use of deep learning pedagogy is effective in increasing students' AI literacy. Initially, students developed a more realistic view of AI, understanding it as a tool with specific limitations rather than "magic." As the teaching model matured through successive cycles, students showed significant improvement in understanding contextual concepts like AI personalization and, most notably, in developing ethical awareness. The study demonstrates that a refined, hands-on pedagogical approach, integrating practical activities with critical reflection on concepts like fairness and privacy, is crucial for successfully fostering comprehensive AI literacy in primary education. These results also indicate that teacher-education programs should include explicit training on facilitating inquiry-based activities, moderating age-appropriate discussions of fairness, privacy, and societal impact, and using reflective cycles to refine practice. School leaders are encouraged to allocate sufficient time, resources, and professional-learning opportunities for teachers to conduct repeated design-enact-reflection cycles, while policymakers can reinforce adoption by providing open educational resources, establishing clear assessment standards for AI-literacy competencies, and promoting collaboration between schools, universities, and industry to keep instructional practices current. Future research should replicate AMAIL with larger and more culturally diverse samples to test generalizability, trace learners longitudinally to examine the durability and transfer of AI-literacy gains, adapt the framework for both younger and older age groups, and develop sensitive instruments that capture incremental growth in students' ethical reasoning; additional studies might also explore links between early AI-literacy education and later academic or career choices, and continue updating pedagogical design principles so that AI-literacy programs remain aligned with evolving technological and societal contexts.

AUTHOR CONTRIBUTIONS

The original draft of the manuscript was written by SCR. The methodology was developed by SCR. Formal analysis was performed by SCR. Investigation and implementation were carried out by SCR. Resources and literature review were provided by JPD. Writing, review, and editing were performed by JPD and SCR. Visualization was produced by SCR. Project administration and correspondence were handled by SCR. All authors have read and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

ACKNOWLEDGEMENTS

This study was supported financially by Satya Wacana Christian University, the United Board for Christian Higher Education in Asia, and the Education University of Hong Kong. Their generous assistance is gratefully acknowledged.

REFERENCES

Almatrafi, O., Johri, A., & Lee, H. (2024). A systematic review of AI literacy conceptualization, constructs, and implementation and assessment efforts (2019–2023). *Computers and Education Open*, *6*, 100173. https://doi.org/10.1016/j.caeo.2024.100173

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher*, *41*(1), 16–25. https://doi.org/10.3102/0013189X11428813

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences*, *13*(1), 1–14. https://doi.org/10.1207/s15327809jls1301_1

Biggs, J. B. (1987). *Student approaches to learning and studying. research monograph.* ERIC.

Chung, K., Kim, S., Jang, Y., Choi, S., & Kim, H. (2025). Developing an AI literacy diagnostic tool for elementary school students. *Education and Information Technologies*, *30*(1), 1013–1044. https://doi.org/10.1007/s10639-024-13097-w

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge University Press. https://doi.org/10.1017/CBO9780511802843

diSessa, A. A. (2022). A history of conceptual change research: Threads and fault lines. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (3rd ed., pp. 114–133). Cambridge University Press; Cambridge Core. https://doi.org/10.1017/9781108888295.008

Donaldson, J. P., Han, A., Yan, S., Lee, S., & Kao, S. (2024). Learning experience network analysis for design-based research. *Information and Learning Sciences*, *125*(1/2), 22–43. https://doi.org/10.1108/ILS-03-2023-0026

Epstein, Z., Hertzmann, A., Investigators of Human Creativity, Akten, M., Farid, H., Fjeld, J., Frank, M. R., Groh, M., Herman, L., & Leach, N. (2023). Art and the science of generative AI. *Science*, *380*(6650), 1110–1111. https://doi.org/10.1126/science.adh4451

Gong, C., Jing, C., Chen, X., Pun, C. M., Huang, G., Saha, A., Nieuwoudt, M., Li, H.-X., Hu, Y., & Wang, S. (2023). Generative AI for brain image computing and brain network computing: A review. *Frontiers in Neuroscience*, *17*, 1203104. https://doi.org/10.3389/fnins.2023.1203104

Google. (2024a, January 4). *Teachable machine*. https://teachablemachine.withgoogle.com/

Google. (2024b, January 15). *AutoDraw*. https://www.autodraw.com/

Google. (2024c, July 5). *Quick, Draw!* https://quickdraw.withgoogle.com/

Heyder, T., & Posegga, O. (2021). *Extending the foundations of AI literacy*. 9. https://aisel.aisnet.org/icis2021/is_future_work/is_future_work/9

Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., Puntambekar, S., & Ryan, M. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by DesignTM into practice. *Journal of the Learning Sciences*, *12*(4), 495–547. https://doi.org/10.1207/S15327809JLS1204_2

Liu, X., & Zhong, B. (2024). A systematic review on how educators teach AI in K-12 education. *Educational Research Review*, *45*, 100642. https://doi.org/10.1016/j.edurev.2024.100642

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3313831.3376727

Ma, M., Ng, D. T. K., Liu, Z., & Wong, G. K. W. (2025). Fostering responsible AI literacy: A systematic review of K-12 AI ethics education. *Computers and Education: Artificial Intelligence*, *8*, 100422. https://doi.org/10.1016/j.caeai.2025.100422

McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, *12*(2), 153–157. https://doi.org/10.1007/BF02295996

Moon, H., Go, H., Lee, Y., & Kim, S.-W. (2024). Investigating factors in artificial intelligence literacy for Korean elementary school students. *International Journal on Advanced Science,* *Engineering and Information Technology*, *14*(4), 1226–1232. Scopus. https://doi.org/10.18517/ijaseit.14.4.16998

Nasir, N. S., Lee, C. D., Pea, R., & McKinney de Royston, M. (2021). Rethinking learning: What the interdisciplinary science tells us. *Educational Researcher*, *50*(8), 557–565. https://doi.org/10.3102/0013189X211047251

Ojeda-Bazaran, M.-J., Chamba-Eras, L., Coronel-Romero, E., Labanda-Jaramillo, M., Conde-Zhingre, L., Irene-Robalino, D., Fierro-Saltos, W., Orellana-Malla, A., Romero-Flores, M., Cueva-Alvarado, G., Chamba-Eras, I., & Erreyes-Pinzon, D. (2021). *State of the art of teaching-learning of artificial intelligence at early ages*. 165–172. Scopus. https://doi.org/10.1109/LACLO54177.2021.00024

Ottenbreit-Leftwich, A., Glazewski, K., Hmelo-Silver, C., Jantaraweragul, K., Jeon, M., Chakraburty, S., Scribner, A., Lee, S., Mott, B., & Lester, J. (2023). *Is elementary AI education possible? 2*, 1364. Scopus. https://doi.org/10.1145/3545947.3576308

Papert, S., & Harel, I. (1991). Situating constructionism. In S. Papert & I. Harel (Eds.), *Constructionism* (pp. 1–11). Basic Books.

Pusat Kurikulum dan Pembelajaran. (2025). *Pembelajaran mendalam—Transformasi pembelajaran menuju pendidikan bermutu untuk semua*. Kementrian Pendidikan Dasar dan Menengah Republik Indonesia.

Relmasira, S. C., Donaldson, J. P., & Lai, Y. C. (2024). Toward a theory-grounded associative model of AI literacy. *Proceedings of the 18th International Conference of the Learning Sciences - ICLS 2024*, 2057–2058. https://doi.org/10.22318/icls2024.559089

Relmasira, S. C., Lai, Y. C., & Donaldson, J. P. (2023). Fostering AI literacy in elementary science, technology, engineering, art, and mathematics (STEAM) education in the age of generative AI. *Sustainability*, *15*(18). https://doi.org/10.3390/su151813595

Sawyer, R. K. (2022). An introduction to the learning sciences. *The Cambridge Handbook of the Learning Sciences*, 1–24. https://doi.org/10.1017/9781108888295.002

Taber, K. S. (2018). The use of cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Touretzky, D., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019). Envisioning AI for K-12: What should every child know about AI? *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*, 9795–9799. https://doi.org/10.1609/aaai.v33i01.33019795

Trigwell, K., & and Prosser, M. (1996). Changing approaches to teaching: A relational perspective. *Studies in Higher Education*, *21*(3), 275–284. https://doi.org/10.1080/03075079612331381211

Voulgari, I., Zammit, M., Stouraitis, E., Liapis, A., & Yannakakis, G. (2021). *Learn to machine learn: Designing a game based approach for teaching machine learning to primary and secondary education students*. 593–598. Scopus. https://doi.org/10.1145/3459990.3465176

Wu, D., Chen, M., Chen, X., & Liu, X. (2024). Analyzing K-12 AI education: A large language model study of classroom instruction on learning theories, pedagogy, tools, and AI literacy. *Computers and Education: Artificial Intelligence*, *7*, 100295. https://doi.org/10.1016/j.caeai.2024.100295

Yang, W. (2022). Artificial Intelligence education for young children: Why, what, and how in curriculum design and implementation. *Computers and Education: Artificial Intelligence*, *3*. Scopus. https://doi.org/10.1016/j.caeai.2022.100061

Yim, I. H. Y. (2024). Artificial intelligence literacy in primary education: An arts-based approach to overcoming age and gender barriers. *Computers and Education: Artificial Intelligence*, *7*. Scopus. https://doi.org/10.1016/j.caeai.2024.100321

Yim, I. H. Y., & Su, J. (2025). Artificial intelligence literacy education in primary schools: A review. *International Journal of Technology and Design Education*. https://doi.org/10.1007/s10798-025-09979-w

Zhang, H., Lee, I., Ali, S., DiPaola, D., Cheng, Y., & Breazeal, C. (2022). Integrating ethics and career futures with technical learning to promote AI literacy for middle school students: An exploratory study. *International Journal of Artificial Intelligence in Education*, 1–35. https://doi.org/10.1007/s40593-022-00293-3