El Midad : Jurnal Jurusan PGMI Vol. 18 No. 2 Agustus 2026
p-ISSN 2087-8389 e-ISSN 2656-4289 p. 492-509
website : http://journal.uinmataram.ac.id/index.php/elmidad email : elmidadpgmi@uinmataram.ac.id

CODING AND ARTIFICIAL INTELLIGENCE IN ELEMENTARY EDUCATION: A SYSTEMATIC REVIEW AND EVIDENCE MAP (2021–2026)

[1]Siswi Yulfani, [2]Lisa Virdinarti Putra
[12]Master of Education Management Study Program, Ngudi Waluyo University
siswiyulfani1@gmail.com; lisavirdinartiputra@gmail.com

ABSTRACT

Coding and artificial intelligence (AI) are increasingly introduced in primary education, yet evidence remains fragmented across computational thinking, AI literacy, digital literacy, ethics, and classroom implementation. This systematic literature review and evidence mapping synthesises primary-school studies published between 2021 and 18 July 2026. A verified screening log was developed through structured keyword searching, publisher and DOI verification, and backwards–forward citation checking. Eligibility was limited to empirical studies involving primary-school students, teachers, or schools and directly examining coding, AI learning, AI literacy, or classroom implementation relevant to pedagogical deep learning. Twenty-three studies were included. The evidence shows a shift from perception-oriented research toward concrete instructional designs, including block-based machine learning, unplugged–plugged sequences, collaborative inquiry, digital storytelling, AIoT, physical computing, analogy-based pedagogy, and debugging. Computational thinking and AI literacy were the most frequently reported outcomes. Stronger learning experiences shared three features: visible AI mechanisms, authentic problem contexts, and structured reflection. Evidence concerning long-term effects, equity, standardised assessment, and explicit deep-learning frameworks remains limited. The synthesis proposes a mechanism linking teacher readiness and infrastructure with pedagogical processes and student outcomes. These findings imply that primary-school implementation should prioritise teacher readiness, age-appropriate scaffolding, meaningful problem solving, and reflection rather than technology use alone, providing direction for curriculum design, teacher development, and educational policy.

Keywords: Coding; Artificial Intelligence; Deep Learning; Primary School; Computational Thinking; AI Literacy

INTRODUCTION

The development of artificial intelligence (AI) has expanded the forms of literacy required from primary education. AI literacy involves not only the ability to use technology but also an understanding of data, system mechanisms, outputs, limitations, and responsible use. (Page et al., 2021). Primary-school students may also enter learning with incomplete conceptions of AI, which strengthens the need for age-appropriate instruction. (Mertala et al., 2022). Coding supports this process by allowing students to organise instructions, test outputs, identify errors, and revise computational solutions. Such activities provide practical opportunities to develop computational thinking while making computational processes more visible to learners. (Moon et al., 2021).

The relevance of this topic has increased in Indonesia following the introduction of Coding and Artificial Intelligence into national education policy (Kemendikdasmen, 2025a). Its implementation at the primary level requires learning activities that are appropriate to students' developmental characteristics and school conditions (Kemendikdasmen, 2025b). This policy direction creates a need to examine international evidence on effective forms of coding and AI learning, reported outcomes, and pedagogical approaches that may inform implementation in Indonesian primary schools.

International research shows a shift from examining students' perceptions and readiness toward more concrete AI learning interventions. Motivation, self-efficacy, and perceived usefulness have been associated with students' intention to learn AI. (Lin et al., 2021). Students' initial conceptions also influence how AI concepts need to be introduced at the primary level. (Rethlefsen et al., 2021). More recent interventions involve neural-network construction, block-based machine learning, physical computing, and AIoT. (Shamir & Levin, 2021). Other studies employ collaborative inquiry, digital storytelling, analogy-based instruction, and unplugged–plugged sequences. (Ruan et al., 2025). These different designs and outcomes make cross-study synthesis necessary to identify broader patterns and limitations.

Previous reviews have mapped AI literacy, pedagogy, tools, assessment, and implementation challenges in primary education. (Yim & Su, 2025). A more specific gap remains in integrating coding, AI learning, and pedagogical deep learning within a single synthesis focused on primary schools. Existing studies show meaningful and contextual applications of AI, but these elements have not been consistently examined as interconnected deep-learning mechanisms. (Ng et al., 2022). The term deep learning in this review therefore refers to a pedagogical approach involving understanding, application, meaningful context, reflection, and active engagement rather than solely to a neural-network algorithm. The novelty of this review lies in mapping how coding and AI implementation, learning outcomes, pedagogical mechanisms, and enabling conditions collectively relate to deep learning in primary education.

The study addresses five research questions: RQ1, what are the publication trends and methodological characteristics of studies published from 2021 to July 18, 2026? RQ2: What forms of coding and AI implementation are used in primary schools? RQ3, what learning outcomes are reported? RQ4, how are the findings related to pedagogical deep learning? RQ5, what enabling factors, barriers, strength of evidence, and research gaps remain? The focus is limited to primary education because younger learners require age-appropriate representation, scaffolding, and learning experiences that differ from those at higher educational levels. (Ottenbreit-Leftwich et al., 2023). The review contributes by integrating evidence on coding, AI literacy, computational thinking, and deep-learning processes to support curriculum development, classroom practice, teacher preparation, and educational policy.

RESEARCH METHODS

Study Design and Protocol

This study employed a Systematic Literature Review (SLR) combined with evidence mapping and thematic synthesis. Study identification and selection were reported according to PRISMA 2020, while the documentation of the search process followed PRISMA-S principles. Evidence mapping was used to describe publication trends, research contexts, study designs, implementation strategies, learning outcomes, and evidence gaps. Thematic synthesis was then applied to identify recurring patterns and relate the findings to pedagogical deep learning while considering differences in the strength of evidence across research designs.

Table 1. Protocol SLR dan evidence mapping

Components	Conditions
Publication range	January 1, 2021–July 18, 2026. This period was selected to capture recent developments in coding and AI education in primary schools, with July 18, 2026 as the final search date.
Population/context	Elementary/primary/elementary students, elementary teachers, or elementary education units
Focus of the intervention	Coding, coding-based computational thinking, AI/KA education, AI literacy, machine learning education, or the implementation of KKA policies in elementary schools
Verification sources	Publisher pages and DOI/Crossref metadata were used to verify bibliographic information, while DOAJ, KCI, institutional indexes, and backwards–forward citation searches were used as supplementary verification and retrieval sources.
Language	Indonesian or English
Document type	Empirical scientific journal articles/proceedings with verifiable metadata and findings
Synthesis units	Primary study; Review articles are used only as a search and comparison source

Search Strategy

The search combined English and Indonesian terms. The core query was ("primary school" OR "elementary school" OR "primary education") AND (coding OR programming OR "computational thinking" OR "artificial intelligence" OR "AI literacy" OR "machine learning education") AND (learning OR teaching OR pedagogy OR curriculum), supplemented by equivalent Indonesian terms. The term deep learning was included only when referring to a pedagogical approach, while studies using it solely as a computational technique were excluded. Candidate records were verified through publisher pages, DOI/Crossref metadata, supplementary indexes, and backwards–forward citation checking. The final search was completed on July 18, 2026.

Only records entered into the verified screening log were counted. After title and DOI normalisation, 35 unique records were identified. Seven records were excluded during title and abstract screening, leaving 28 articles for full-text assessment. Five additional articles were excluded after full-text review because the primary-school sample was not separable, the educational level was outside the review scope, or the study did not provide relevant evidence on coding or AI learning. Consequently, 23 primary studies were included in the final synthesis.

Inclusion and Exclusion Criteria

Eligibility was determined according to publication period, educational level, empirical design, and direct relevance to coding or AI education. Findings were considered relevant when they addressed classroom implementation, learning processes, computational thinking, AI literacy, digital literacy, ethical awareness, problem solving, motivation, or related educational outcomes. Pedagogical deep learning referred to learning involving understanding, application, contextualization, reflection, or active engagement, whereas studies using *deep learning* solely as a computational algorithm were excluded.

Table 2. Inclusion and exclusion criteria

Aspects	Operationalization
Inclusions	Published 2021–2026; Elementary/Primary/Elementary context; empirical research; discussing coding/KA/literacy KA/ML education/CT related to coding activities; findings related to classroom implementation, learning processes, computational thinking, AI literacy, digital literacy, ethical awareness, problem solving, motivation, or other relevant educational outcomes.
Exclusion	JUNIOR HIGH SCHOOL, HIGH SCHOOL, VOCATIONAL SCHOOL, college; cross-level samples that do not separate SD data; deep learning as an algorithm without an educational focus; purely conceptual articles; service/training that does not provide evidence of elementary school implementation; reviews that are not primary studies.
Age limit rules	Mixed-age studies are only accepted when the entire range is still in elementary school age, or the researcher explicitly states the elementary/primary context.
Term rules	Technical deep learning, machine learning, and pedagogical deep learning are separated from the screening to prevent mixing of concepts.

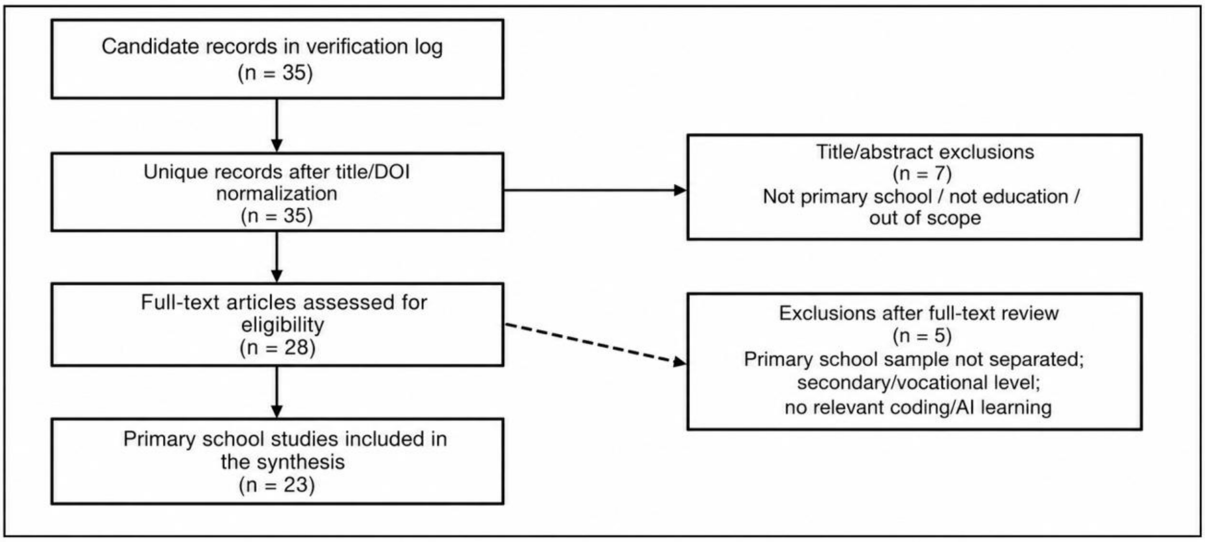


Figure 1. Flow of study identification, screening, eligibility assessment, and inclusion.

Extraction, Coding, and Weighting of Evidence

Each study was extracted into a matrix containing publication year, country or region, target group, research design, implementation strategy, learning outcomes, relevance to pedagogical deep learning, and relative strength of evidence. Coding was conducted at the study level, allowing one article to receive more than one strategy or outcome code. Related codes were compared across studies to identify broader patterns concerning implementation, outcomes, pedagogical mechanisms, and evidence gaps. Frequencies were interpreted as the density of themes within the corpus rather than as measures of intervention effect.

Table 3. Synthesis codebook and proof-weighting rules

Code	Operational definition
CT	Decomposition, patterns, abstractions, algorithms, debugging, or computational thinking measurement
AI literacy	Understanding of AI/ML concepts, data, models, usage, implementation, or evaluation of AI outputs
Digital literacy	Ability to use and assess digital technologies that go beyond device operation
Ethical/critical	Bias, privacy, risk, responsibility, system boundaries, or social reflection
Meaningful	Authentic issues, everyday experiences, contextual projects, or cross-subject applications
Reflection	Evaluation, debugging, reasoning discussion, output assessment, or reflection on process/risk
Evidence Weight A	Comparative, experimental, or quasi-experimental studies providing stronger evidence of differences or changes in outcomes.
Evidence Weight B	Mixed-methods, large surveys, design-based studies, or multi-source cases providing moderate empirical support for patterns or mechanisms.
Evidence Weight C	Small-scale exploratory or single-group studies providing preliminary evidence of feasibility, perceptions, or implementation.

The A–C categories indicate relative evidentiary weight based on research design and inferential capacity rather than a formal risk-of-bias assessment.

RESULTS AND DISCUSSION

The discussion is anchored in sociocultural theory, which emphasises learning through active participation, interaction, and scaffolding within meaningful activities. (Vygotsky, 1978). This perspective is aligned with pedagogical deep learning because students are expected not only to receive technical information but also to construct understanding, apply knowledge, solve problems, and reflect on learning processes. Coding and AI are therefore interpreted as pedagogical tools whose value depends on how they support these learning processes.

Corpus Characteristics

The final corpus consists of 23 primary studies. Publication frequency did not increase linearly, but the evidence shows a clear thematic shift from studies of perceptions and motivation toward more concrete instructional interventions. Six studies were published in 2021, three studies each year from 2022 to 2025, and five studies by July 18, 2026. The 2026 total represents partial-year data and should therefore not be interpreted as a complete annual trend.

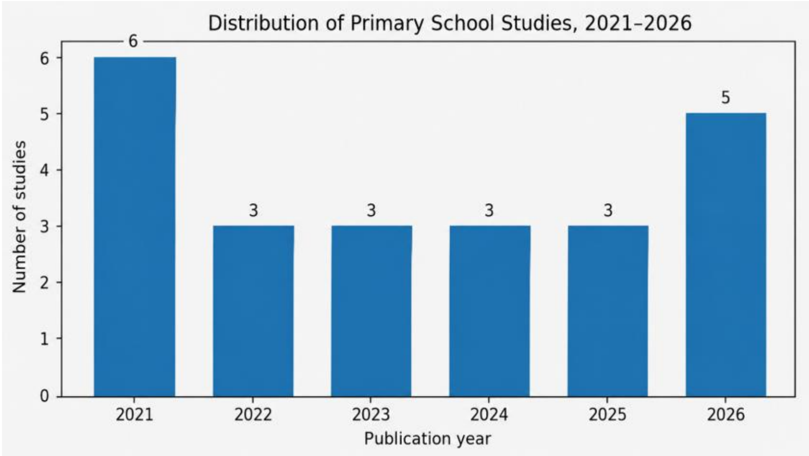


Figure 2. Distribution of 23 primary studies of elementary school.

Geographically, the evidence is concentrated in East and Southeast Asia, with additional studies from the United States and Europe. Indonesian studies become more visible in the later period, particularly as coding and AI enter primary-school policy and classroom implementation. This concentration limits broad generalisation because infrastructure, curriculum, teacher readiness, and school conditions vary across national settings.

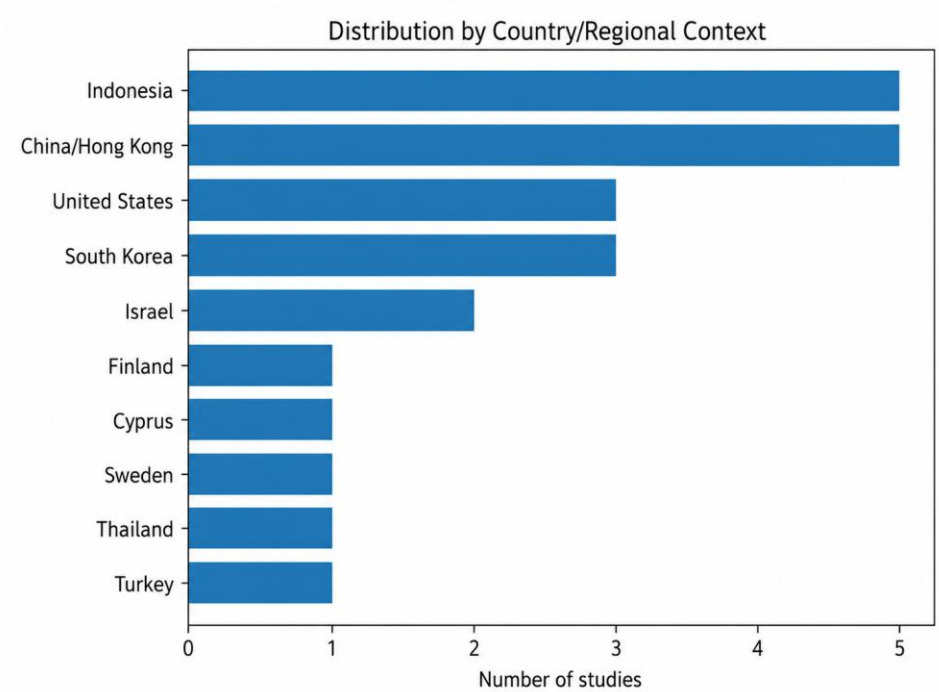
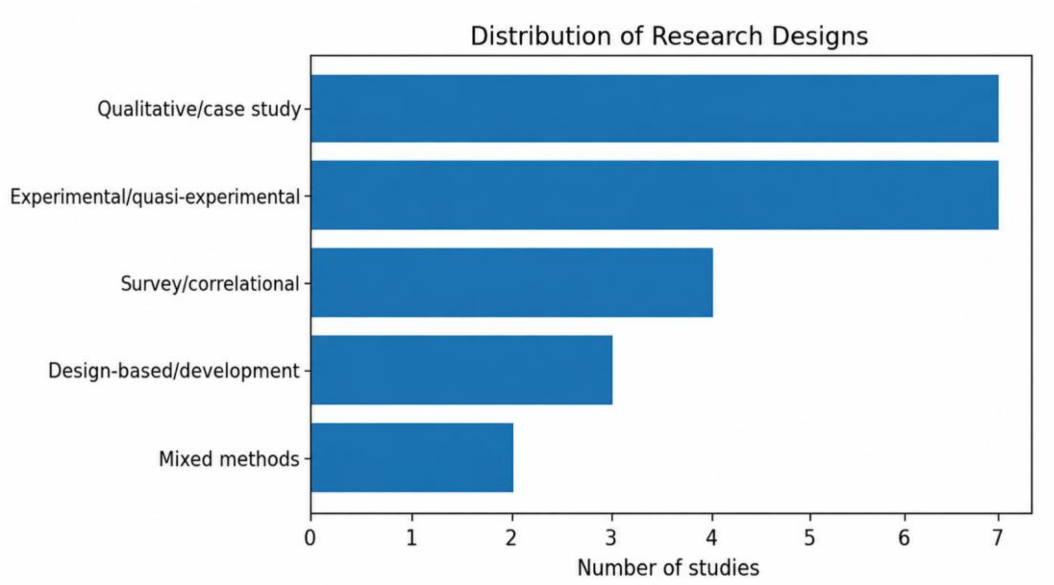


Figure 3. Distribution of the context of the primary study country/region.

The methodological pattern also shifts from surveys, case studies, and design-based research toward experimental and quasi-experimental designs. Seven comparative or quasi-experimental studies provide stronger evidence of outcome changes, whereas qualitative and mixed-methods studies clarify implementation processes, teacher readiness, and pedagogical



mechanisms. The combination is important because evidence of effectiveness does not by itself explain why an intervention works differently across classroom contexts. A combination of the two sets of evidence is necessary because the question of whether an intervention works is different from the question of why the intervention works or fails. **Figure 4. Distribution of research designs in corpus.**

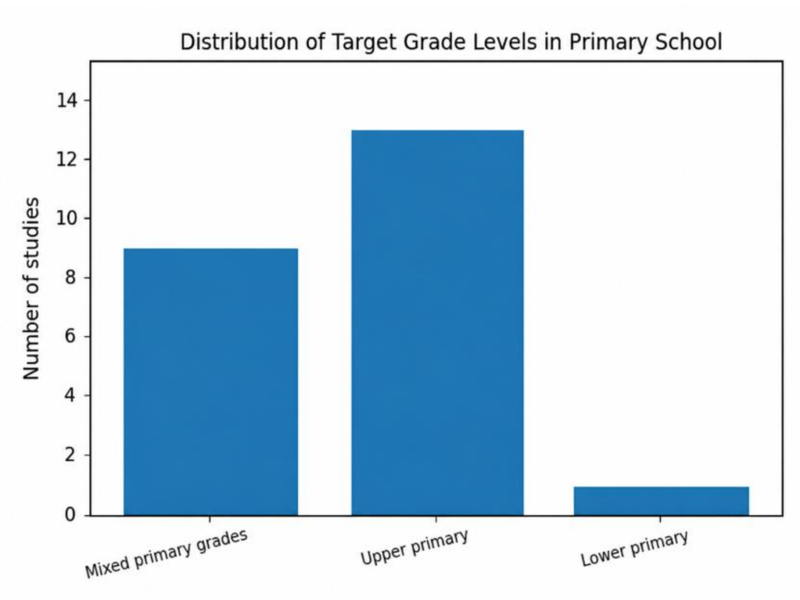


Figure 5. Tier target distribution

Primary Study Matrix Analysed

The following matrix summarises the 23 primary studies included in the synthesis. All studies met the primary-school eligibility criteria, while cross-level studies without separable elementary-school data were excluded to maintain consistency in developmental and educational context.

Table 4. Summary of 23 primary school studies 2021–2026

No.	Study	Context	Target Group	Design	Relevant Findings
1	Lin et al. (2021)	China/Hong Kong	Mixed primary school grades	Survey/correlational	A structural model of motivation to learn AI highlights the roles of perceived usefulness, self-efficacy, and learning readiness as key foundations for AI adoption in primary education.
2	Chai et al., (2021)	China/Hong Kong	Mixed primary school grades	Survey/correlational	A survey of 682 primary school students showed that self-efficacy, readiness, perceived social usefulness, and AI literacy were associated with students' intention to learn AI.
3	Shamir & Levin (2021)	Israel	Upper primary	Qualitative/case study	The construction of neural networks by 12-year-old students showed that model-building activities helped reveal AI mechanisms that had previously been perceived as a black box.
4	Lee et al., (2021)	United States	Upper primary	Design-based/developmental	PrimaryAI integrates games, collaborative inquiry, and life science problems to introduce AI concepts to students aged 8–11 years.
5	Druga & Ko (2021)	United States	Mixed primary school grades	Mixed methods	A four-week program involving training and coding of intelligent programs changed how children understood machine intelligence and human–AI relationships.
6	Moon et al. (2021)	South Korea	Upper primary	Experimental/quasi-experimental	Block-based machine learning with numerical data was developed and tested to strengthen primary school students' computational thinking.
7	Shamir & Levin (2022)	Israel	Upper primary	Experimental/quasi-experimental	Machine learning through learning-by-design and learning-by-teaching expanded computational thinking toward data-driven ways of thinking.
8	Ng et al. (2022)	China/Hong Kong	Mixed primary school grades	Mixed methods	Three months of digital story writing involving 82 students encouraged the application of AI knowledge to formulate solutions related to real-world problems.
9	Mertala et al. (2022)	Finland	Upper primary	Survey/correlational	A qualitative survey of 195 Grade 5–6 students identified diverse initial conceptions of AI, including anthropomorphism and limited understanding of the role of data.
10	Ottenbreit-Leftwich et al. (2023)	United States	Upper primary	Qualitative/case study	A study involving Grade 4–5 students and teachers mapped everyday experiences, conceptions, and ethical issues as entry points for the co-design of a primary-level AI curriculum.
11	Relmasira et al., (2023)	Indonesia	Upper primary	Design-based/developmental	Design-based research in primary schools integrated unplugged activities, Teachable Machine, smart speakers, and deepfake-related issues to develop AI literacy and ethical reflection.
12	Jaratsaeng	Thailand	Lower	Design-	An unplugged coding package

	et al. (2023)		primary	based/developmental	supported by AI-assisted tools was implemented with 18 Grade 2 students using a one-group posttest design to support computational thinking.
13	Dai et al., (2024)	China/Hong Kong	Upper primary	Experimental/quasi-experimental	A human–AI analogy approach in upper primary education outperformed direct instruction in AI knowledge, skills, and ethical awareness.
14	Hong & Kim (2025)	South Korea	Mixed primary school grades	Experimental/quasi-experimental	An AIoT program demonstrated the potential of integrating AI with real-life problems to strengthen primary school students' digital literacy and AI literacy.
15	Yoo & Kim (2024)	South Korea	Upper primary	Experimental/quasi-experimental	Students participating in AI-integrated physical computing showed greater improvement in computational thinking than those engaged in physical computing without AI.
16	Walan (2025)	Sweden	Upper primary	Qualitative/case study	A case study of students aged 11–12 years mapped their cognitive and affective perceptions of AI before and after instruction, including views on its benefits and risks.
17	Kalemkuş & Kalemkuş (2025)	Türkiye	Mixed primary school grades	Qualitative/case study	Analysis of metaphors and drawings from 262 Grade 3–4 students revealed diverse conceptions of AI and highlighted the importance of beginning instruction from students' existing representations.
18	Ruan et al. (2025)	China/Hong Kong	Upper primary	Experimental/quasi-experimental	A repeated quasi-experiment involving 56 Grade 4 students compared unplugged–plugged instruction with plugged-only instruction. The clearest advantage appeared in AI perceptions, while effects on self-efficacy were inconsistent.
19	Aisyah et al. (2026)	Indonesia	Upper primary	Experimental/quasi-experimental	A Coding and Artificial Intelligence program for Grade 5 students was tested in relation to computational thinking skills, providing initial evidence from the Indonesian implementation context.
20	Mardhiyah et al. (2026)	Indonesia	Mixed primary school grades	Qualitative/case study	A study involving teachers and students in two primary schools showed that contextual activities, problem solving, active participation, and reflection supported decomposition, pattern recognition, abstraction, and algorithmic thinking.
21	Susanto et al. (2026)	Indonesia	Mixed primary school grades	Qualitative/case study	A collective case study identified four models of Coding and Artificial Intelligence implementation in primary schools and showed that teachers' TPACK readiness and school leadership influenced the depth of implementation.
22	Kesek et al. (2026)	Indonesia	Upper primary	Qualitative/case study	Implementation of a Grade 5 Coding and Artificial Intelligence module emphasised decomposition, algorithms, debugging, and reflection. Pedagogical depth was considered more influential

El Midad : Jurnal Jurusan PGMI Vol. 18 No. 2 Agustus 2026
p-ISSN 2087-8389 e-ISSN 2656-4289 p. 492-509
website : http://journal.uinmataram.ac.id/index.php/elmidad email : elmidadpgmi@uinmataram.ac.id

					than merely using digital applications.
23	Kasinidou et al., (2026)	Cyprus	Mixed primary school grades	Survey/correlational	A survey of 233 students across seven schools mapped AI perceptions, gaps in prior knowledge, and their relationship with digital skills as a basis for designing responsible AI education.

Evolution of Coding and AI Implementation in Primary Schools

The strategy map shows that plugged, machine-learning, and AI-based tools are the most frequently reported forms of implementation. Digital technology, however, is not necessarily the most appropriate starting point for every learner. Unplugged activities can function as a conceptual bridge before students interact with more complex digital systems (Ruan et al., 2025). This sequence allows abstract ideas such as classification, patterns, decisions, and algorithms to be introduced through more concrete experiences.

Greater pedagogical depth appears when students move from using AI toward constructing, testing, or explaining how a system works. Neural-network construction makes otherwise hidden AI mechanisms more observable to learners (Shamir & Levin, 2021). Digital storytelling connects AI knowledge with authentic problems and solution development (Ng et al., 2022). Physical computing also links computational reasoning with tangible system behaviour (Kim & Kim, 2024). These findings support the theoretical view that learning becomes deeper through active construction and meaningful application rather than through technological complexity alone.

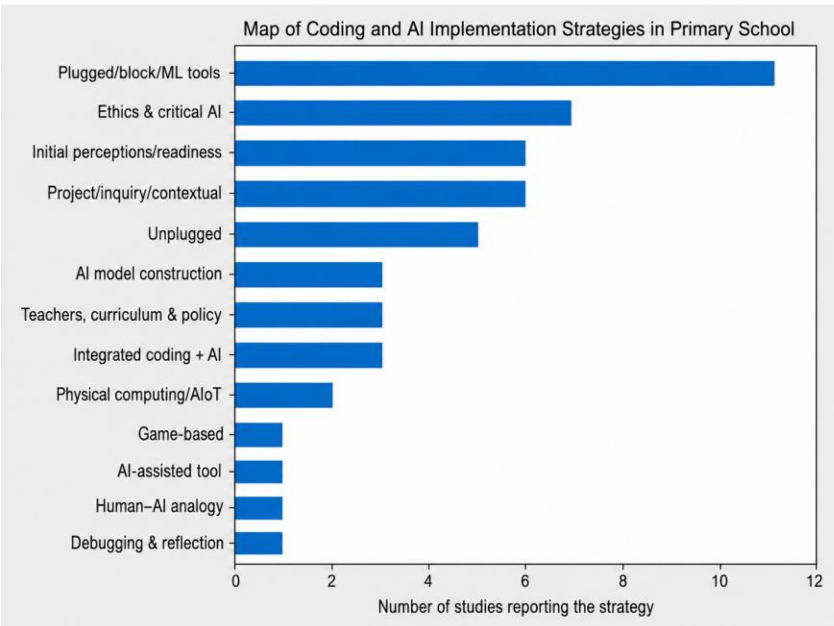


Figure 6. Frequency of implementation strategies coded from 23 studies

Learning Outcomes: From Computational Thinking to Critical AI Literacy

AI literacy is the most broadly reported because it includes understanding AI concepts, use, application, and evaluation, while computational thinking is especially prominent in coding, block-based machine learning, physical computing, and debugging activities. Block-based machine learning has produced evidence of computational-thinking development among elementary-school students (Moon et al., 2021). AI-integrated physical computing also showed stronger computational-thinking improvement than physical computing without AI integration (Yoo & Kim, 2024). Indonesian evidence further indicates that decomposition, pattern recognition, algorithmic thinking, and debugging emerge as important learning processes during coding and AI activities (Mardhiyah et al., 2026).

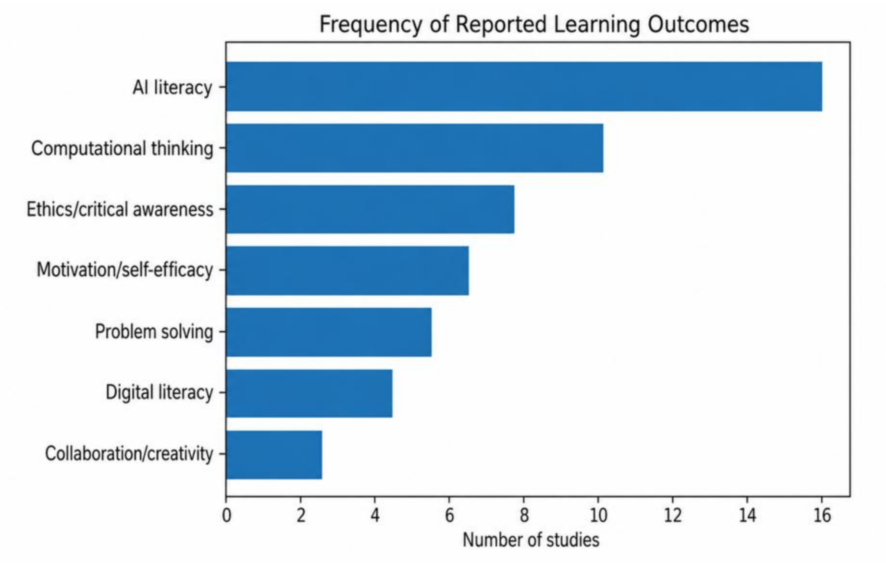


Figure 7. Frequency of reported achievements

Ethical awareness, problem solving, motivation, and digital literacy also appear across the corpus, but their assessment is less consistent. This difference suggests that the evidence for computational thinking and AI literacy is currently more established than the evidence for broader affective, ethical, and collaborative outcomes. AI literacy also requires a balance between technical understanding and critical judgment. Primary-school students may hold incomplete conceptions of how AI works before formal instruction begins (Mertala et al., 2022). Students also express both positive and negative perceptions of AI after encountering its benefits and possible risks (Walan, 2025). Initial diagnostic assessment is therefore important before instruction progresses toward application, evaluation, and ethical reflection.

El Midad : Jurnal Jurusan PGMI | Vol. 18 No. 2 Agustus 2026
p-ISSN 2087-8389 e-ISSN 2656-4289 | p. 492-509
website : http://journal.uinmataram.ac.id/index.php/elmidad | email : elmidadpgmi@uinmataram.ac.id

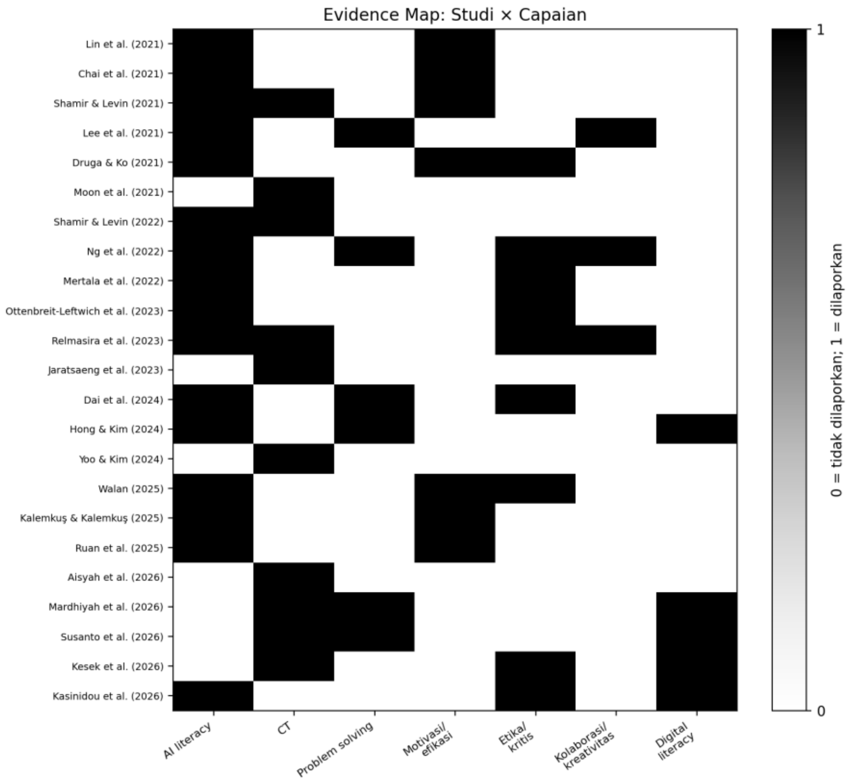


Figure 8. Evidence map of the study × achievements.

Coding and AI as Vehicles for Pedagogical Deep Learning

The thematic synthesis shows the strongest relationship with understanding and application, whereas explicit reflection appears less frequently. This pattern is important because pedagogical deep learning requires students to move beyond executing technical procedures toward examining reasons, errors, alternatives, and consequences. Debugging provides a concrete opportunity for students to reconsider both the process and the quality of a computational solution. (Kesek et al., 2026). Ethical evaluation extends reflection from technical accuracy toward the social consequences of AI use (Dai et al., 2024).

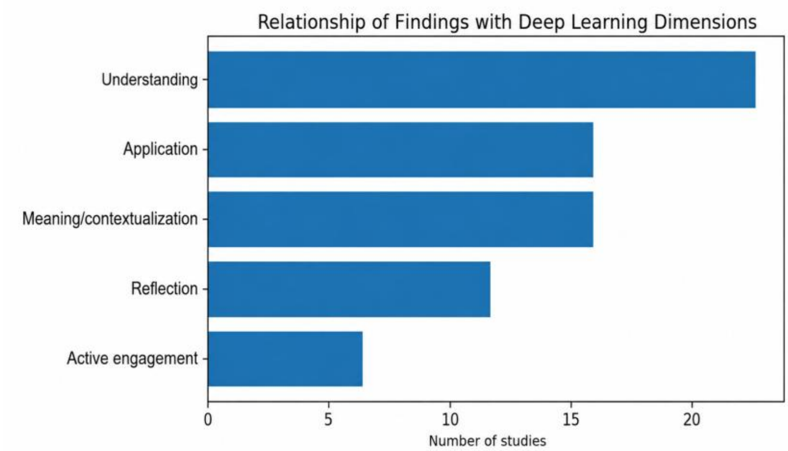


Figure 9. The relationship of empirical findings with the dimension of deep learning.

El Midad : Jurnal Jurusan PGMI Vol. 18 No. 2 Agustus 2026
p-ISSN 2087-8389 e-ISSN 2656-4289 p. 492-509
website : http://journal.uinmataram.ac.id/index.php/elmidad email : elmidadpgmi@uinmataram.ac.id

Three mechanisms consistently connect coding and AI with deeper learning: making AI mechanisms visible, situating learning in authentic problems, and providing structured opportunities for revision and reflection. Analogy-based instruction helps make abstract AI processes more understandable to primary-school students (Dai et al., 2024). Contextual AI projects enable students to apply knowledge rather than only reproduce technical steps (Hong & Kim, 2024). Structured reflection through debugging and evaluation allows students to examine how and why a solution works. Together, these mechanisms explain why the pedagogical design of an activity is more important than the sophistication of the technology itself.

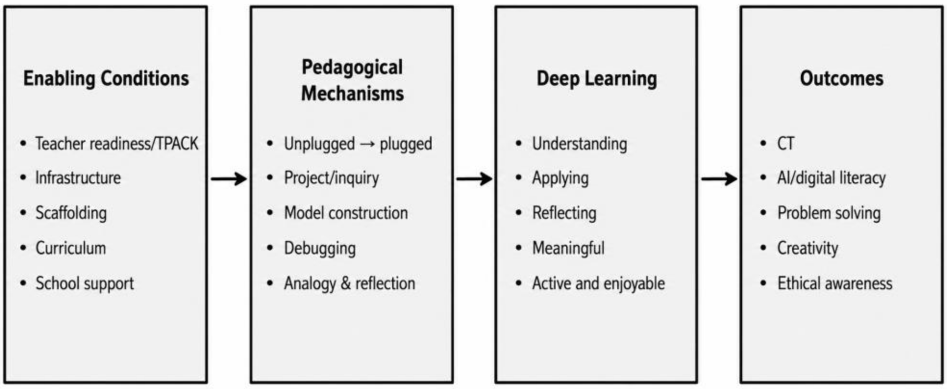


Figure 10. Synthesis model of coding and AI implementation mechanisms for pedagogical deep learning in primary education.

The main contribution of this synthesis is therefore not a claim that one technology is superior, but the identification of a common pedagogical mechanism across heterogeneous interventions. Coding and AI appear to support deep learning most plausibly when students can observe a mechanism, apply it to a meaningful problem, and reconsider the resulting process or output.

Relative Weight of Evidence

Relative evidence weighting prevents findings from different research designs from being interpreted as equally conclusive. Weight A studies provide comparatively stronger evidence of outcome changes, while Weight B studies are more informative for identifying patterns and implementation mechanisms. Weight C studies remain useful for understanding feasibility and early implementation but provide a weaker basis for broad effectiveness claims.

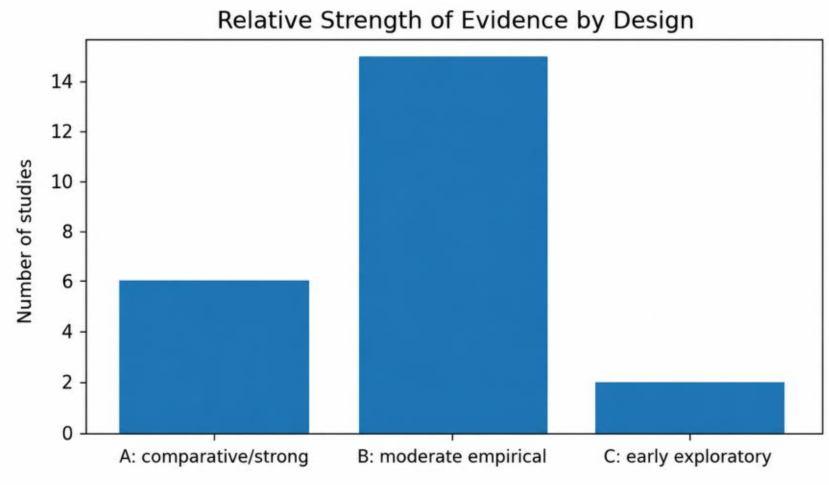


Figure 11. The distribution of relative strength of evidence based on the research design.

The most defensible interpretation is that coding and AI are associated with stronger computational thinking and AI literacy when learning is active, age-appropriate, contextual, and supported by scaffolding. Universal causal claims remain premature because study duration, instruments, samples, and school contexts vary considerably. Importantly, positive results are not uniform across all outcomes; unplugged–plugged instruction improved some aspects of AI learning while self-efficacy gains remained inconsistent (Ruan et al., 2025). This contrasting evidence challenges the assumption that the introduction of AI technology automatically improves every dimension of student learning.

Enabling Factors and Obstacles

Teacher readiness emerges as a central link between curriculum expectations and students' actual learning experiences. Curriculum co-design can help align AI instruction with students' prior conceptions and teachers' classroom knowledge (Ottenbreit-Leftwich et al., 2023). Differences in teacher TPACK and school leadership also influence the depth of implementation across primary schools (Susanto et al., 2026). Infrastructure remains relevant, but limited devices do not necessarily prevent conceptual learning because unplugged activities can introduce sequences, patterns, classification, and algorithms before digital practice. The evidence therefore suggests that pedagogical readiness and appropriate sequencing are at least as important as access to sophisticated technology.

El Midad : Jurnal Jurusan PGMI | Vol. 18 No. 2 Agustus 2026
p-ISSN 2087-8389 e-ISSN 2656-4289 | p. 492-509
website : http://journal.uinmataram.ac.id/index.php/elmidad | email : elmidadpgmi@uinmataram.ac.id

Evidence Gap and Research Agenda

The main evidence gap concerns the depth and sustainability of research rather than the absence of topics. Long-term follow-up is rare, while evidence concerning equitable access, socioeconomic differences, and the effects of infrastructure remains limited. Explicit use of pedagogical deep-learning frameworks is also relatively recent, even though earlier studies contain meaningful, contextual, or reflective learning characteristics.

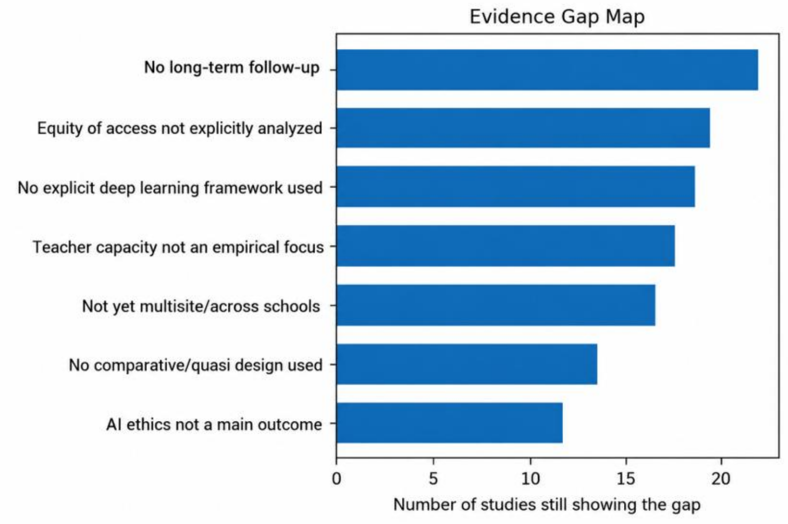


Figure 12. Evidence gap map based on corpus coding

Future research should prioritise multisite designs, longer follow-up periods, and more consistent instruments for computational thinking and AI literacy. Comparative studies are also needed to examine unplugged–plugged sequencing, intervention duration, debugging, teacher scaffolding, and ethical learning outcomes. Indonesian implementation particularly requires cross-regional evidence because schools operate with different levels of digital infrastructure and teacher readiness.

Synthesis of Answers to Research Questions

Table 5. Synthesis of Answers to the Research Questions

Questions	Synthesis answer
RQ1: Publication trends and study characteristics	Twenty-three primary studies published between 2021 and July 18, 2026 were included. The evidence shifts from perception and motivation studies toward experimental, design-based, and implementation-oriented research.
RQ2: Forms of implementation	The main strategies include plugged or block-based AI tools, contextual projects and inquiry, unplugged activities, ethical-critical learning, physical computing or AIoT, and curriculum implementation supported by teacher readiness.
RQ3: Learning outcomes	AI literacy and computational thinking are the most frequently reported

	outcomes, followed by problem solving, motivation or self-efficacy, ethical awareness, creativity or collaboration, and digital literacy.
RQ4: Relationship with pedagogical deep learning	The strongest links appear in understanding and application, while reflection becomes visible through debugging, output evaluation, ethical discussion, and comparison of human and AI decisions.
RQ5: Enabling factors and evidence gaps	Teacher readiness, scaffolding, infrastructure, authentic problems, and activity sequencing influence implementation. Major gaps concern longitudinal evidence, equitable access, standardised assessment, and explicit testing of pedagogical deep-learning frameworks.

CONCLUSION

The synthesis of 23 primary studies shows that coding and AI education has developed toward more constructive, contextual, and reflective learning experiences. Computational thinking and AI literacy are the most consistently reported outcomes, while evidence for ethical awareness, problem solving, digital literacy, motivation, and creativity remains more varied. The relationship with pedagogical deep learning becomes strongest when students can understand how a system works, apply concepts to authentic problems, test or revise solutions, and reflect on technological outputs and consequences.

The findings indicate that educational value depends more on pedagogical design than on technological complexity. Coding and AI are most likely to support deep learning when implementation combines age-appropriate scaffolding, visible system mechanisms, meaningful problems, and structured reflection. Evidence is still insufficient to support universal effectiveness claims because study duration, instruments, research designs, and school conditions remain heterogeneous. The synthesis therefore contributes an implementation mechanism that links pedagogical readiness, meaningful application, and reflection with student learning outcomes in primary education.

SUGGESTIONS

Schools should prioritise pedagogical and teacher readiness before technological provision. Teacher development should integrate TPACK, computational thinking, AI literacy, assessment, ethics, and deep-learning strategies, while primary-school modules should balance understanding, application, and reflection with unplugged alternatives for schools with limited digital access. Future research should employ multisite designs, comparison groups, longer follow-up periods, and more consistent assessment instruments to strengthen evidence of effectiveness and equitable implementation.

REFERENCE

Aisyah, A. S., Abidin, Z., Wicaksono, V. D., & Penehafo, A. E. (2026). The influence of the coding and artificial intelligence program on fifth graders' computational thinking skills. *Sekolah Dasar: Kajian Teori dan Praktik Pendidikan*, *35*(1), 195–212. https://doi.org/10.17977/um009v35i12026p195-212

Chai, C. S., Lin, P.-Y., Jong, M. S.-Y., Dai, Y., Chiu, T. K. F., & Qin, J. (2021). Perceptions of and behavioural intentions towards learning artificial intelligence in primary school students. *Educational Technology & Society*, *24*(3), 89–101.

Dai, Y., Lin, Z., Liu, A., Dai, D., & Wang, W. (2024). Effect of an analogy-based approach of artificial intelligence pedagogy in upper primary schools. *Journal of Educational Computing Research*, *61*(8), 1695–1722. https://doi.org/10.1177/07356331231201342

Druga, S., & Ko, A. J. (2021). How do children's perceptions of machine intelligence change when training and coding smart programs? *Proceedings of the Interaction Design and Children Conference*, 49–61. https://doi.org/10.1145/3459990.3460712

Hong, J., & Kim, K. (2025). Impact of AIoT education program on digital and AI literacy of elementary school students. *Education and Information Technologies*, *30*, 107–130. https://doi.org/10.1007/s10639-024-12758-0

Jaratsaeng, K., Kanjug, I., & Jarungsirawat, N. (2023). Unplugged coding learning package and A.I. assisted learning tool to support computational thinking for elementary school students. *ILAI Letters on Institutional Research*, *3*. https://doi.org/10.52731/lir.v003.149

Kalemkuş, J., & Kalemkuş, F. (2025). Primary school students' perceptions of artificial intelligence: Metaphor and drawing analysis. *European Journal of Education*, *60*, e70007. https://doi.org/10.1111/ejed.70007

Kasinidou, M., Kleanthous, S., & Otterbacher, J. (2026). What do children think about AI?: Insights and educational implications from primary school students' perceptions of AI. *Proceedings of the AAAI Conference on Artificial Intelligence*, *40*(47), 40608–40616. https://doi.org/10.1609/aaai.v40i47.41507

Kemendikdasmen. (2025a). *Keputusan Menteri Pendidikan Dasar dan Menengah Republik Indonesia Nomor 127/P/2025 tentang Pedoman Implementasi Koding dan Kecerdasan Artifisial pada Pendidikan Anak Usia Dini, Jenjang Pendidikan Dasar, dan Jenjang Pendidikan Menengah*. Kementerian Pendidikan Dasar dan Menengah Republik Indonesia.

Kemendikdasmen. (2025b). *Naskah akademik pembelajaran koding dan kecerdasan artifisial pada pendidikan dasar dan menengah*. Kementerian Pendidikan Dasar dan Menengah Republik Indonesia.

Kesek, J., Bachri, B. S., & Kholidya, C. F. (2026). Analisis implementasi modul ajar Kurikulum Merdeka pada mata pelajaran Koding dan Kecerdasan Artifisial (KKA) di kelas 5 SD Negeri 4 Mogolaing. *Jurnal Studi Pendidikan Dasar*, *4*(1), 15–27. https://doi.org/10.54180/jsped.2026.4.1.15-27

Kim, Y. S., & Kim, Y. S. (2024). The effect of physical computing programming education integrating artificial intelligence on computational thinking ability of elementary school students. *Journal of the Korea Society of Computer and Information*, *29*(3), 227–235. https://doi.org/10.9708/jksci.2024.29.03.227

Lee, S., Mott, B., Ottenbreit-Leftwich, A., Scribner, A., Taylor, S., Park, K., Rowe, J., Glazewski, K., Hmelo-Silver, C. E., & Lester, J. (2021). AI-infused collaborative inquiry in upper elementary school: A game-based learning approach. *Proceedings of the AAAI Conference on Artificial Intelligence*, *35*(17), 15591–15599. https://doi.org/10.1609/aaai.v35i17.17836

Lin, P.-Y., Chai, C.-S., Jong, M. S.-Y., Dai, Y., Guo, Y., & Qin, J. (2021). Modelling the structural relationship among primary students' motivation to learn artificial intelligence. *Computers and Education: Artificial Intelligence*, *2*, 100006. https://doi.org/10.1016/j.caeai.2020.100006

Mardhiyah, A., Hadi, M. S., & Misriandi. (2026). Deep learning-based coding and artificial intelligence for computational thinking in elementary schools. *Academia Open*, *11*(1). https://doi.org/10.21070/acopen.11.2026.13221

Mertala, P., Fagerlund, J., & Calderon, O. (2022). Finnish 5th and 6th grade students' pre-instructional conceptions of artificial intelligence (AI) and their implications for AI literacy education. *Computers and Education: Artificial Intelligence*, *3*, 100095. https://doi.org/10.1016/j.caeai.2022.100095

Moon, W., Lee, J., Kim, B., Seo, Y., Kim, J., Oh, J., Kim, Y., & Kim, J. (2021). Effect of block-based machine learning education using numerical data on computational thinking of elementary school students. *Journal of the Korean Association of Information Education*, *25*(2), 367–375. https://doi.org/10.14352/jkaie.2021.25.2.367

Ng, D. T. K., Luo, W., Chan, H. M. Y., & Chu, S. K. W. (2022). Using digital story writing as a pedagogy to develop AI literacy among primary students. *Computers and Education: Artificial Intelligence*, *3*, 100054. https://doi.org/10.1016/j.caeai.2022.100054

Ottenbreit-Leftwich, A., Glazewski, K. D., Jeon, M., Jantaraweragul, K., Hmelo-Silver, C. E., Scribner, A., Lee, S., Mott, B., & Lester, J. (2023). Lessons learned for AI education with elementary students and teachers. *International Journal of Artificial Intelligence in Education*, *33*(2), 267–289. https://doi.org/10.1007/s40593-022-00304-3

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71

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), 13595. https://doi.org/10.3390/su151813595

Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., Koffel, J. B., & Group, P.-S. (2021). PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Systematic Reviews*, *10*, 39. https://doi.org/10.1186/s13643-020-01542-z

Ruan, J., Ma, H., Sun, J., Chu, H., Jing, B., & Zhang, Y. (2025). Developing artificial intelligence literacy through mixed unplugged and plugged-in activities in primary education. *SAGE Open*, *15*(4), 1–15. https://doi.org/10.1177/21582440251336708

Shamir, G., & Levin, I. (2021). Neural network construction practices in elementary school. *KI - Künstliche Intelligenz*, *35*, 181–189. https://doi.org/10.1007/s13218-021-00729-3

Shamir, G., & Levin, I. (2022). Teaching machine learning in elementary school. *International Journal of Child-Computer Interaction*, *31*, 100415. https://doi.org/10.1016/j.ijcci.2021.100415

Susanto, E., Sugiyanti, W., Arifin, M., Fathonah, S., & Nurkolis. (2026). Enacting coding and artificial intelligence education policy in elementary schools. *Journal of Innovation and Research in Primary Education*, *5*(1), 1616–1628. https://doi.org/10.56916/jirpe.v5i1.3135

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (ed.)). Harvard University Press.

Walan, S. (2025). Primary school students' perceptions of artificial intelligence – for good or bad. *International Journal of Technology and Design Education*, *35*, 25–40. https://doi.org/10.1007/s10798-024-09898-2

Yim, I. H. Y., & Su, J. (2025). Artificial intelligence literacy education in primary schools: A review. *International Journal of Technology and Design Education*, *35*, 2175–2204. https://doi.org/10.1007/s10798-025-09979-w