Jurnal Inovasi dan Riset Pendidikan Dasar 1 (3) September 2026: 183-196 E-ISSN: 3125-0157
DOI: https://doi.org/10.64421/jirpd.v1i3.130

The Effect of a Pembelajaran Mendalam (Mindful, Meaningful, and Joyful Learning) Coding and Artificial Intelligence Module on Elementary School Students' Computational Thinking

Andayani Andayani [1,*], Suryo Prabowo [2], Rina Kartika Sari [3], Dimas Aditya Pratama [3]

[1] Department of Teacher Professional Education Universitas Terbuka, Indonesia
[2] Department of Educational Technology Study Program, Universitas Terbuka, Indonesia
[3] Department of Elementary Teacher Education, Universitas Terbuka, Indonesia

*Corresponding author: andayani@ecampus.ut.ac.id

To cite this article: Andayani, A., Prabowo, S., Sari, R. K., & Pratama, D. A. (2026). The Effect of a Pembelajaran Mendalam (Mindful, Meaningful, and Joyful Learning) Coding and Artificial Intelligence Module on Elementary School Students' Computational Thinking. *Jurnal Inovasi dan Riset Pendidikan Dasar, 1*(3), 183-196. https://doi.org/10.64421/jirpd.v1i3.130

Informasi Artikel

Diterima : 13-10-2026
Direvisi : 24-10-2026
Diterima : 26-10-2026
Dipublikasi: 01-11-2026

Abstrak

Mulai tahun ajaran 2025/2026, Koding dan Kecerdasan Artifisial (KKA) diperkenalkan sebagai mata pelajaran pilihan mulai kelas V sekolah dasar. Penelitian ini membandingkan modul KKA berbasis Pembelajaran Mendalam (mindful, meaningful, dan joyful), yaitu pendekatan pedagogis dan bukan deep learning dalam pembelajaran mesin, dengan pembelajaran berbasis aplikasi. Studi kuasi-eksperimen nonequivalent control group melibatkan 65 siswa kelas V di dua sekolah dasar Tangerang Selatan (eksperimen n = 33; kontrol n = 32). Perlakuan berlangsung enam pertemuan, masing-masing 70 menit, menggabungkan aktivitas unplugged, Scratch, klasifikasi gambar sederhana, dan refleksi. Data diperoleh melalui tes berpikir komputasional, angket pengalaman belajar, dan observasi. Rerata posttest eksperimen (76,88) melebihi kontrol (65,40), t(63) = 4,73, p < 0,001, d = 1,17; N-gain masing-masing 0,58 dan 0,37. Secara deskriptif, peningkatan indikator terbesar terjadi pada dekomposisi (0,62) dan terkecil pada abstraksi (0,47) di kelompok eksperimen. Dalam sampel ini, penggunaan modul berkaitan dengan peningkatan berpikir komputasional yang lebih besar daripada pembelajaran berbasis aplikasi; keterbatasan desain dua sekolah menghalangi atribusi kausal yang tegas.

Kata kunci: Berpikir Komputasional; Kecerdasan Artifisial; Koding; Pembelajaran Mendalam; Sekolah Dasar; Unplugged.

Abstract

Starting in the 2025/2026 academic year, Coding and Artificial Intelligence (Koding dan Kecerdasan Artifisial, KKA) was introduced as an elective subject from grade five in Indonesian elementary schools. This quasi-experimental study compared a KKA module based on Pembelajaran Mendalam (mindful, meaningful, and joyful learning)—a pedagogical approach distinct from machine-learning-based deep learning—with application-only instruction. Sixty-five fifth-grade students from two schools in South Tangerang participated (experimental n = 33; control n = 32). Across six 70-minute meetings, the module integrated unplugged activities, Scratch, a simple image-classification tool, and reflection. Computational thinking tests, a learning-experience questionnaire, and observations supplied the data. Posttest mean scores were 76.88 (experimental) and 65.40 (control), t(63) = 4.73, p < .001, d = 1.17; corresponding N-gains were 0.58 and 0.37. Descriptively, the largest indicator gain in the experimental group was decomposition (0.62), while the smallest was abstraction (0.47). In this two-school sample, the module was associated with greater computational-thinking gains than application-only instruction; the design does not establish which component produced the difference.

Keywords: Computational Thinking; Coding; Pembelajaran Mendalam; Elementary School; Scratch; Bebras.

Copyright ©2026 Andayani Andayani, Suryo Prabowo, Rina Kartika Sari, Dimas Aditya Pratama
Jurnal Inovasi dan Riset Pendidikan Dasar is licensed under a
Creative Commons Attribution-ShareAlike 4.0 International License

1. INTRODUCTION

The rapid development of artificial intelligence (AI), particularly generative AI, requires education systems to prepare students not only as users of technology but also as thinkers who understand how it works and what its consequences are. In response, Indonesia's Ministry of Primary and Secondary Education introduced Coding and Artificial Intelligence (Koding dan Kecerdasan Artifisial, KKA) as an elective subject beginning in the 2025/2026 academic year, starting from the fifth grade of elementary school (Kemendikdasmen, 2025a). The core of this subject is the development of computational thinking, digital literacy, and a basic, ethical understanding of AI. The implementation framework is set out in Ministerial Decree No. 127/P/2025 and the curriculum amendment in Regulation No. 13/2025; phase C learning activities encompass structured problem solving and introductory coding and AI concepts.

Computational thinking refers to the ability to solve problems using ways of thinking common in computer science, including decomposition, pattern recognition, abstraction, and algorithm design (Wing, 2006; Grover & Pea, 2013). A meta-analysis shows that learning to program yields cognitive benefits that transfer to other thinking skills, such as reasoning and creativity (Scherer et al., 2019). However, implementing KKA in elementary schools faces several challenges: limited computer devices in many schools, teachers who mostly do not have a background in informatics, and the risk of lessons that focus on introducing applications without building ways of thinking.

Alongside this policy, the government is also promoting the Pembelajaran Mendalam approach (mindful, meaningful, and joyful learning; not neural-network deep learning), which emphasises mindful, meaningful, and joyful learning experiences (Kemendikdasmen, 2025b). This approach suits the needs of KKA instruction in elementary schools, because abstract computing concepts need to be connected to real problems in students' surroundings and experienced directly. Integrating Deep Learning with local wisdom in elementary schools has shown promising results (Andayani, 2025b), but its application to the KKA subject has rarely been studied.

One strategy relevant to the conditions of Indonesian schools is to combine unplugged (computer-free) and plugged (device-based) activities. Unplugged activities have been shown to develop computational thinking effectively in the early years of primary education (del Olmo-Muñoz et al., 2020), while plugged activities using visual programming languages allow students to test and debug algorithms directly (Brennan & Resnick, 2012; Relkin et al., 2021). Meanwhile, the digital literacy readiness of elementary teachers and pre-service teachers still varies (Prabowo et al., 2023; Prabowo & Andayani, 2025), so teachers need structured learning modules that are easy to implement. Application-only demonstrations may encourage students to reproduce steps without first articulating a solution strategy or reflecting on errors. A combined approach can instead make algorithmic reasoning visible before students translate it into blocks; however, the present comparison tests a package of instructional features rather than any one component in isolation. A recent meta-analysis of 49 studies found promising overall effects of unplugged approaches on K–12 computational thinking, while emphasising variability across contexts (Chen et al., 2023). Evidence from

Indonesian elementary-school literature reviews also indicates growing interest in unplugged coding, but does not by itself establish the comparative effectiveness of this particular KKA module (Kasiono et al., 2025).

Based on this background, this study aimed to examine the effect of a Deep Learning-based KKA module combining unplugged and plugged activities on the computational thinking of fifth-grade elementary school students. The research questions were: (1) Is there a significant difference in computational thinking between students taught with the Deep Learning-based KKA module and students taught with application-only KKA instruction? (2) Which computational thinking indicators are most affected? and (3) How do students experience the learning in terms of the mindful, meaningful, and joyful principles? Students' reported learning experiences are examined alongside test performance because engagement and reflection are proposed pedagogical processes, not substitutes for achievement outcomes. The study addresses a context-specific evidence gap; its two-school design does not establish that similar interventions have not been studied elsewhere in Indonesia.

2. LITERATURE REVIEW

This section reviews computational thinking, introductory AI literacy, and the instructional rationale for a combined unplugged–plugged module under the Indonesian Pembelajaran Mendalam approach. The comparison concerns a multicomponent teaching design, not a direct test of individual module elements.

2.1. Computational Thinking and AI Literacy in Elementary School Students

Wing (2006) introduced computational thinking as a fundamental skill for everyone, not only computer scientists. Grover and Pea (2013) described its components as abstraction and pattern generalisation, systematic information processing, problem decomposition, and algorithm design. Brennan and Resnick (2012) added that children's computational thinking develops through computational concepts, practices, and perspectives built while creating meaningful projects. As AI has advanced, Touretzky et al. (2019) proposed five big ideas of AI to be introduced from early schooling, including the idea that computers learn from data, while Long and Magerko (2020) defined AI literacy as the ability to understand, use, and critically evaluate AI. At the elementary level, AI literacy can be introduced through simple activities, such as sorting data to train a classification model. Bebras is an international informatics challenge whose problem-based tasks are often used to foster and probe computational thinking; adaptation of this task format requires independent validation in the target age group (Dagienė & Stupurienė, 2016).

2.2. Deep Learning in KKA Instruction

In Indonesian education policy, Deep Learning is understood as an approach that honours students through mindful, meaningful, and joyful learning experiences, organised in a cycle of understanding, applying, and reflecting (Kemendikdasmen, 2025b). In KKA instruction, the mindful principle is realised through students' reflection on their own problem-solving strategies; the meaningful principle through real

problems, such as scheduling class duties or sorting waste; and the joyful principle through games, collaboration, and projects. Differentiated instruction is also important because students vary in their readiness to use technology (Andayani & Iasha, 2023). Bers (2018) argues that coding for children should be positioned as a playground that offers room for exploration and creativity, not merely syntax drills. In principle, reflection prompts may support metacognitive monitoring: pupils revisit instructions, detect mismatches between planned and observed outputs, and revise algorithms. This hypothesised pathway is consistent with the emphasis on iterative testing and debugging in Brennan and Resnick (2012), but it was not independently tested as a mediator here. The conceptual expectation is that explicit problem decomposition, practical enactment, and subsequent reflection may jointly support computational thinking. Accordingly, the directional hypothesis is that the experimental group will show higher posttest computational-thinking scores than the application-only group; indicator-specific results and learning experience are exploratory. Evidence syntheses suggest that unplugged activities are potentially beneficial, but their effects vary by context and study design (Chen et al., 2023).

3. METHOD

This study used a quantitative approach with a quasi-experimental nonequivalent control group design (Shadish et al., 2002). The experimental group learned with a Deep Learning-based KKA module combining unplugged and plugged activities, while the control group learned KKA through the approach commonly used in schools, namely teacher demonstration followed by direct practice in a visual programming application. Both groups covered the same learning outcomes and content and took a pretest and a posttest. As participants were members of pre-existing school groups and were not individually randomised, unmeasured school-, teacher-, and selection-related factors may contribute to the observed contrasts. Furthermore, unplugged tasks, reflection journals, and the distribution of activity time differed simultaneously; the effects of these components cannot be isolated.

3.1. Participants

Participants were 65 fifth-grade students from two public elementary schools in South Tangerang City who had chosen KKA as an elective subject in the even semester of the 2025/2026 academic year. Both schools had the same accreditation and computer laboratories with a ratio of one device per two students. The experimental group consisted of 33 students (17 girls, 16 boys) and the control group of 32 students (16 girls, 16 boys). None of the students had received formal programming instruction before. KKA enrolment was elective, so students' choice to participate may limit generalisability. Group membership is confounded with school in the available description; the available dataset and narrative do not establish whether teaching personnel and prior experience were equivalent across conditions.

3.2. Module and Treatment

The module was developed by the research team through collaboration between the Teacher Professional Education and Educational Technology programs, and was validated by two informatics

content experts, one elementary pedagogy expert, and one media expert, who rated it as highly feasible. It comprised six meetings (6 × 70 minutes) on: (1) decomposing everyday problems, (2) patterns and loops, (3) algorithms and instruction sequences, (4) conditionals, (5) how machines learn from data, and (6) a mini project. Each meeting followed the understand–apply–reflect cycle: students began with an unplugged activity (for example, a human-robot game on a grid floor or sorting picture cards to train a classification model), then applied it in a block-based visual programming application (Scratch) and a simple web-based image classification tool, and closed with a reflection journal. The final project asked students to design an interactive animation about waste sorting at school. The control group studied the same topics for the same duration, but all activities were carried out directly in the application following teacher demonstration, without unplugged activities or reflection journals. Because validation scores, the experts' scoring rubric and cut-offs, and detailed lesson plans are not provided in the supplied manuscript, the reported "highly feasible" judgement cannot be independently verified from the present article. These materials should be supplied as supplementary files when available. The existing description indicates that both groups completed the fifth-session AI classification topic, but the precise tools and exposure time used by the control group require author confirmation.

3.3. Instruments and Data Collection

The main instrument was a 20-item computational thinking test in multiple-choice and short-answer formats, adapted from Bebras challenge task types and contextualised for elementary students. The test measured four indicators: decomposition (5 items), pattern recognition (5 items), abstraction (4 items), and algorithms (6 items). Content validity was reviewed by three experts, and a trial with 36 students outside the sample yielded a KR-20 reliability coefficient of 0.81. Supporting instruments included a 15-item, four-point Likert learning experience questionnaire measuring the three Deep Learning principles (mindful, meaningful, and joyful) and an implementation observation sheet completed by two observers at every meeting. Bebras-inspired tasks were used as a task-format reference rather than evidence that the adapted items retain the original assessment's psychometric properties. The manuscript does not specify whether the identical items were reused on both occasions; possible practice effects therefore cannot be ruled out. The reported KR-20 = 0.81 concerns the total score; indicator-level reliability estimates for the short subscales and separate reliability and validity evidence for the experience questionnaire are unavailable in the supplied materials. The criteria defining "very good" questionnaire results also require documentation. The task format was informed by the Bebras model (Dagienė & Stupurienė, 2016).

3.4. Data Analysis

Data were analysed using descriptive statistics, assumption tests (Shapiro-Wilk normality test and Levene's homogeneity test), and an independent samples t-test at a 0.05 significance level. Effect size was calculated using Cohen's d. Gains were analysed with the normalised gain (N-gain) (Hake, 1998), categorised as high ($g \geq 0.70$), medium ($0.30 \leq g < 0.70$), and low ($g < 0.30$). Questionnaire data were analysed descriptively as percentages, and students' reflection journals were analysed thematically to enrich the

interpretation of the results. These tests provide unadjusted between-group comparisons. Because the design includes pretest measurement and intact groups, ANCOVA with pretest as a covariate would be preferable if the original student-level data and its model assumptions were available. No ANCOVA estimates are asserted here in the absence of that dataset. Indicator-level differences and questionnaire ratings are treated as descriptive and exploratory rather than as separately tested intervention effects. The manuscript does not report the number of reflection journals, a documented coding protocol, or inter-coder checks; quotations are therefore illustrative rather than evidence of a systematically validated qualitative analysis.

4. RESULTS AND DISCUSSION

The results are presented in five parts: (1) implementation fidelity, (2) description of computational thinking, (3) assumption and hypothesis testing, (4) gains by computational thinking indicator, and (5) students' learning experience. The discussion then interprets these findings in relation to theory and previous research.

4.1. Implementation Fidelity

Observations showed that implementation in the experimental group reached an average of 92.7% across the six meetings, with the lowest score at the first meeting (85.0%) and the highest at the fourth to sixth meetings (95.0%). At the first meeting, the unplugged activity took longer than planned because students were very enthusiastic, so the reflection session was shortened. The control group reached an average of 94.2%. Observers in both groups noted problems mainly with the internet connection when using the web-based image classification tool at the fifth meeting. Observed implementation was generally high in both groups, but fidelity by itself cannot rule out school, teacher, selection, or novelty effects or establish a causal attribution. Internet connectivity difficulties during the fifth meeting may also have affected engagement with the AI classification activity.

4.2. Description of Computational Thinking

Descriptive statistics of the computational thinking test scores for the experimental and control groups at pretest and posttest are presented in Table 1.

Table 1. Descriptive statistics of computational thinking scores

Group	Test	Min	Max	Mean	SD	N
Experimental	Pretest	25	70	45.47	10.12	33
Experimental	Posttest	55	95	76.88	9.05	33
Control	Pretest	25	65	44.92	9.86	32
Control	Posttest	40	90	65.40	10.47	32

Source: Research data (2026)

Table 1 shows similar observed pretest means (45.47 and 44.92; difference 0.55 points), although a nonsignificant baseline test does not prove equivalence of the two intact groups. Following instruction, the experimental mean increased by 31.41 points to 76.88 and the control mean by 20.48 points to 65.40. The experimental group's posttest standard deviation (9.05) was lower than its pretest standard deviation (10.12). These aggregate descriptive results do not establish improvement specifically among students with low initial achievement, which would require subgroup or individual-level analysis.

4.3. Assumption and Hypothesis Testing

Before hypothesis testing, the data were tested for normality and homogeneity. A summary of the assumption tests is presented in Table 2.

Table 2. Results of normality and homogeneity tests

Data	Group	Shapiro-Wilk (W)	Sig.	Levene (F)	Sig.
Pretest	Experimental	0.960	0.254	0.05	0.824
Pretest	Control	0.957	0.231		
Posttest	Experimental	0.949	0.121	1.37	0.246
Posttest	Control	0.963	0.327		

Source: Research data (2026)

Table 2 shows that all Shapiro-Wilk significance values were greater than 0.05, indicating normally distributed data. Levene's test significance values for the pretest (0.824) and posttest (0.246) were also greater than 0.05, indicating homogeneous variances. Therefore, the independent samples t-test was appropriate. The results are presented in Table 3.

Table 3. Results of the independent samples t-test

Data	Mean Difference	t	df	Sig. (2-tailed)	95% CI	Cohen's d
Pretest	0.55	0.22	63	0.825	[−4.40, 5.50]	0.06
Posttest	11.48	4.73	63	< 0.001	[6.63, 16.33]	1.17
N-gain	0.21	5.13	63	< 0.001	[0.13, 0.29]	1.27

Source: Research data (2026)

The t-test on pretest data showed no significant difference between the groups ($p = 0.825$; $d = 0.06$), which does not establish group equivalence. On the posttest, there was a significant difference with a mean difference of 11.48 points, $t(63) = 4.73$, $p < 0.001$. The 95% confidence interval ranged from 6.63 to 16.33 and did not include zero. The effect size, Cohen's $d = 1.17$, is large, meaning the experimental group's mean was more than one standard deviation above the control group's. The t-test on N-gain scores also showed a significant difference ($t(63) = 5.13$; $p < 0.001$; $d = 1.27$). The observed comparison supports an association between the multicomponent module and higher posttest scores in this sample; causality and the contribution of individual activities remain uncertain.

4.4. Gains by Computational Thinking Indicator

To describe the pattern of changes across computational-thinking components, indicator scores were summarised using N-gain. As no inferential comparison between indicators was conducted, the indicator ranking below is descriptive only. Results are shown in Table 4.

Table 4. Mean scores and N-gain by computational thinking indicator

Indicator	Group	Pretest	Posttest	N-gain	Category
Decomposition	Experimental	48.5	80.6	0.62	Medium
Decomposition	Control	47.9	67.8	0.38	Medium
Pattern recognition	Experimental	46.1	78.2	0.60	Medium
Pattern recognition	Control	45.6	66.3	0.38	Medium
Abstraction	Experimental	41.8	68.9	0.47	Medium
Abstraction	Control	41.2	58.1	0.29	Low
Algorithms	Experimental	45.0	78.5	0.61	Medium
Algorithms	Control	44.8	67.4	0.41	Medium

Source: Research data (2026)

Table 4 shows that the experimental group outperformed the control group on all four indicators. The highest gain was in decomposition (N-gain 0.62). Items for this indicator asked students to break a complex task, such as planning a classroom clean-up, into small, manageable steps. The unplugged activity at the start of each meeting accustomed students to writing out solution steps on paper before assembling code blocks, so decomposition was practised explicitly.

The algorithms indicator reached an N-gain of 0.61 in the experimental group and 0.41 in the control group. The between-group difference for algorithms is 0.20 N-gain points; the smallest descriptive gap is for abstraction (0.18), while decomposition shows a gap of 0.24. Both groups practised sequencing blocks in Scratch, but the study did not statistically test differences among these gaps. The difference lay in the quality of answers: students in the experimental group more often found and fixed sequencing errors (debugging) in short-answer items, while control students tended to use trial and error without tracing the steps. Reflection journals confirmed this, for example: "I checked the steps again one by one, like when I was the robot on the floor."

Pattern recognition showed an N-gain of 0.60 in the experimental group. The meeting on how machines learn from data appears to have strengthened this indicator, as students grouped picture cards by shared features before training a simple classification model. Several students wrote that "the computer can be wrong if there are only a few example pictures," showing an emerging understanding of the role of data in AI. Because internet problems were observed during the fifth session, an influence on practice opportunities or pattern-recognition outcomes cannot be excluded without session-level or student-level data.

Abstraction showed the smallest gain in both groups (0.47 and 0.29). Abstraction requires students

to ignore irrelevant details and form general representations, such as using a variable for a game score. This ability depends on formal reasoning that is still developing in 10–11-year-olds, so it requires longer practice and more explicit scaffolding.

The distribution of individual N-gain categories reinforces these findings. In the experimental group, 6 students (18.2%) were in the high category, 24 (72.7%) in the medium category, and 3 (9.1%) in the low category. In the control group, 2 students (6.3%) were in the high category, 18 (56.3%) in the medium category, and 12 (37.5%) in the low category. The proportion of students with low gains in the control group was more than four times that of the experimental group.

4.5. Students' Learning Experience

The experimental group's learning experience was examined in terms of the three Deep Learning principles. The questionnaire results are summarised in Table 5.

Table 5. Students' learning experience by Deep Learning principle

Principle	Number of Items	Percentage	Category
Mindful	5	82.4%	Very good
Meaningful	5	87.3%	Very good
Joyful	5	91.5%	Very good
Mean	15	87.1%	Very good

Source: Research data (2026)

Table 5 shows that all principles fell in the very good category, with a mean of 87.1%. The joyful principle scored highest (91.5%); the highest-rated item was "I enjoy learning coding because there are games" (94.7%). The meaningful principle scored 87.3%, mainly because students felt the waste-sorting project related to their lives. The mindful principle scored lowest (82.4%). Thematic analysis of the reflection journals showed that in the first two meetings some students wrote only "fun" or "hard," but in later meetings they began describing their strategies and the errors they had fixed. As this questionnaire was administered only to the experimental group, the percentages describe satisfaction with that condition and do not support a comparison of student experience between conditions. The scoring boundaries underlying the "very good" designation should be documented by the authors.

4.6. Discussion

4.6.1. Association Between the Pembelajaran Mendalam-Based KKA Module and Computational Thinking

The findings show that the Deep Learning-based KKA module was associated with higher scores in students' computational thinking with a large effect size (d = 1.17). Notably, the control group also improved meaningfully (N-gain 0.37), indicating that KKA instruction of any form is beneficial. However, the experimental group's advantage shows that how KKA is taught matters greatly. This is consistent with del Olmo-Muñoz et al. (2020), who found that starting with unplugged activities before plugged ones

produced better computational thinking development than plugged activities alone. For context, Chen et al. (2023) estimated an overall Hedges's g of 1.028 for unplugged interventions across a heterogeneous K–12 literature; this is not directly equivalent to the present unadjusted posttest Cohen's d = 1.17, and differences in design preclude straightforward effect-size comparisons.

The understand–apply–reflect cycle of Deep Learning (Kemendikdasmen, 2025b) provided a structure in which computing concepts were experienced before being formalised. Unplugged activities acted as a bridge between concrete experience and symbolic representation in code blocks. This aligns with Brennan and Resnick's (2012) view that children's computational thinking grows through designing, testing, and debugging rather than through conceptual explanation alone. It is also consistent with Scherer et al.'s (2019) meta-analysis showing that well-designed programming instruction yields transferable cognitive benefits.

4.6.2. Differences in Gains across Indicators

The pattern of highest gains in decomposition and lowest in abstraction is consistent with findings that abstraction is the most difficult component of computational thinking for elementary-age children (Grover & Pea, 2013; Relkin et al., 2021). Decomposition and algorithms are easier to practise because they can be visualised as concrete steps, whereas abstraction requires generalisation. Future KKA modules should therefore add activities dedicated to abstraction, such as comparing several games to find a general rule, with gradually faded scaffolding. For example, a future lesson could ask students to compare several Scratch games, identify which details can be ignored, articulate a shared rule, and transfer it to a new game without worked examples. The observed indicator order is descriptive, and the short indicator subscales have no separately reported reliability estimates.

The emerging understanding of the role of data in AI seen in students' journals suggests that big ideas of AI, such as "computers learn from data" (Touretzky et al., 2019), can be introduced to fifth graders through simple activities. This matters because today's students are already exposed to AI-based applications in daily life. AI literacy, including the ability to critically evaluate AI outputs (Long & Magerko, 2020), can begin in elementary school, for example by discussing why a classification model can make mistakes.

4.6.3. Learning Experience and Practical Implications

The high score for the joyful principle supports Bers's (2018) view that coding for children should be framed as a space for play and creation. Meanwhile, the lowest score for the mindful principle shows that students' metacognitive skills need to be cultivated. Teachers should provide specific reflection prompts, such as "Which step was the hardest, and how did you fix it?", so that reflection goes beyond expressing feelings.

In terms of policy implementation, these findings carry three implications. First, unplugged activities may be useful in resource-limited settings, although this possibility must be tested in schools with fewer

devices because both study schools had one device for every two students. Second, most elementary teachers do not have an informatics background and therefore need structured learning modules and continuous digital literacy training (Andayani, 2025a; Prabowo & Andayani, 2025). Third, Teacher Professional Education programs need to equip pre-service elementary teachers with KKA pedagogical competence, including designing Deep Learning-based instruction and understanding the ethical aspects of AI use.

5. CONCLUSION

This two-school quasi-experiment found higher posttest computational-thinking scores in the group using a Pembelajaran Mendalam-based KKA module combining unplugged activities, plugged work and reflection than in the group receiving application-only instruction. Descriptively, decomposition showed the greatest N-gain and abstraction the smallest in the experimental condition; between-indicator differences were not tested inferentially. Students in the experimental group reported positive experiences, especially for joyful learning, but no comparable questionnaire responses were collected for controls. The study did not assess AI literacy as a separate outcome. Limitations include its small and geographically restricted sample, intact-school assignment, possible school and teacher confounding, elective-subject self-selection, simultaneous variation in several intervention components, a six-meeting duration, and potential novelty or researcher-development effects. Further research should use a larger multisite design, pretest-adjusted analyses, independently validated subscales, transparent qualitative coding and follow-up assessment. The module represents a promising instructional option for further evaluation rather than an established generalisable model for all elementary schools.

6. ACKNOWLEDGEMENTS

The authors thank the Institute for Research and Community Service (LPPM) of Universitas Terbuka, as well as the principals, teachers, and fifth-grade students of the partner schools in South Tangerang City for their support and participation in this study.

Author Contributions

Andayani Andayani: conceptualisation, methodology, supervision, writing – original draft. Suryo Prabowo: module design and development, technological resources, formal analysis, writing – review and editing. Rina Kartika Sari: investigation, data curation. Dimas Aditya Pratama: investigation, analysis of reflection journals, project administration.

Funding Statement

This research did not receive any funding.

Ethics Statement

This study obtained permission from the relevant education office and schools. Written informed consent was obtained from students' parents or guardians, and participants' identities were kept confidential. The web-based image classification tool was used without uploading any students' personal data.

Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Conflict of Interest Statement

The authors declare no conflict of interest.

Statement on the Use of *Artificial Intelligence*

The authors used Chatgpt to assist in preparing the initial draft and language editing. All content has been reviewed and verified by the authors, who take full responsibility for it.

REFERENCES

Andayani, A. (2025a). Peningkatan literasi digital dan kreativitas guru sekolah dasar melalui pelatihan augmented reality di Jakarta Timur. Indonesian Journal of Community Service in Education, 1(2), 80–89.

Andayani, A. (2025b). Transformasi pembelajaran di sekolah dasar Kepulauan Seribu: Integrasi deep learning, kearifan lokal, dan nilai karakter. Indonesian Journal of Community Service in Education, 1(1), 22–35.

Andayani, A., & Iasha, V. (2023). Looking at differentiated learning from the lens of underachiever students: Viewpoints, experiences, and impacts. Journal of Social Research, 3(1), 201–208.

Bers, M. U. (2018). Coding as a playground: Programming and computational thinking in the early childhood classroom. Routledge.

Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. American Educational Research Association Annual Meeting. https://scratched.gse.harvard.edu/ct/files/AERA2012.pdf

del Olmo-Muñoz, J., Cózar-Gutiérrez, R., & González-Calero, J. A. (2020). Computational thinking through unplugged activities in early years of primary education. Computers & Education, 150, 103832. https://doi.org/10.1016/j.compedu.2020.103832

Grover, S., & Pea, R. (2013). Computational thinking in K–12: A review of the state of the field. Educational Researcher, 42(1), 38–43. https://doi.org/10.3102/0013189X12463051

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. American Journal of Physics, 66(1), 64–74. https://doi.org/10.1119/1.18809

Kemendikdasmen. (2025a). Keputusan Menteri Pendidikan Dasar dan Menengah Nomor 127/P/2025 tentang pedoman implementasi koding dan kecerdasan artifisial pada pendidikan anak usia dini, jenjang pendidikan dasar, dan jenjang pendidikan menengah. https://www.kemendikdasmen.go.id/pengumuman/13654-kepmendikdasmen-tentang-pedoman-implementasi-koding-dan-kecerdasan-artifisial-pada-paud-dikdasmen

Kemendikdasmen. (2025b). Pembelajaran mendalam: Menuju pendidikan bermutu untuk semua [Deep learning: Towards quality education for all; Academic paper]. Ministry of Primary and Secondary Education of the Republic of Indonesia.

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

Prabowo, S., & Andayani, A. (2025). Augmented reality as catalyst for digital literacy enhancement: Investigating university students' readiness in the era of disruptive educational technologies. International Journal of Education and Learning Studies, 1(1), 37–48.

Prabowo, S., Andayani, A., & Hanafi, H. (2023). Literasi digital dalam pembelajaran: Perspektif alumni PGSD. Jurnal Basicedu, 7(1), 99–105.

Relkin, E., de Ruiter, L. E., & Bers, M. U. (2021). Learning to code and the acquisition of computational thinking by young children. Computers & Education, 169, 104222. https://doi.org/10.1016/j.compedu.2021.104222

Scherer, R., Siddiq, F., & Sánchez Viveros, B. (2019). The cognitive benefits of learning computer programming: A meta-analysis of transfer effects. Journal of Educational Psychology, 111(5), 764–792. https://doi.org/10.1037/edu0000314

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin.

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(1), 9795–9799. https://doi.org/10.1609/aaai.v33i01.33019795

Kasiono, B., Pradhana, G., Huda, N., Utami, S., & Haerussaleh. (2025). Unlocking young minds: The impact of unplugged coding on elementary students' computational thinking. Jurnal Paedagogy, 12(3). https://doi.org/10.33394/jp.v12i3.15650

Dagienė, V., & Stupurienė, G. (2016). Bebras—A sustainable community building model for the concept based learning of informatics and computational thinking. Informatics in Education, 15(1), 25–44. https://doi.org/10.15388/infedu.2016.02

Chen, P., Yang, D., Metwally, A. H. S., Lavonen, J., & Wang, X. (2023). Fostering computational thinking through unplugged activities: A systematic literature review and meta-analysis. International Journal of STEM Education, 10, Article 47. https://doi.org/10.1186/s40594-023-00434-7

Wing, J. M. (2006). Computational thinking. Communications of the ACM, 49(3), 33–35. https://doi.org/10.1145/1118178.1118215