Prisma Sains: Jurnal Pengkajian Ilmu dan Pembelajaran Matematika dan IPA IKIP Mataram
https://e-journal.undikma.ac.id/index.php/prismasains/index
e-mail: prismasains.pkpsm@gmail.com

July 2026. Vol. 14, No. 3
p-ISSN: 2338-4530
e-ISSN: 2540-7899
pp. 1167-1186

Integrating Coding and Project-Based Learning to Foster Elementary Students' Creativity

***Leni Agustina Daulay, Rahmah Nurfitriani**

Faculty Islamic Education, Institut Agama Islam Negeri Takengon, Takengon, Indonesia

*Corresponding Author e-mail: agustina_leni@yahoo.com

Received: March 2026; Revised: May 2026; Published: July 2026

Abstract

Despite growing interest in coding education for fostering 21st-century skills, empirical evidence from resource-constrained, non-metropolitan settings remains limited, particularly when comparing digital and non-digital approaches. This study investigated the comparative effectiveness of plugged and unplugged coding instruction within a Project-Based Learning (PjBL) framework in enhancing creative thinking skills among fifth-grade students. Using a mixed-methods quasi-experimental design, 98 students from two Integrated Islamic Elementary Schools in Central Aceh were assigned to a plugged coding group ($n = 50$) using Scratch 3.0 or an unplugged coding group ($n = 48$) using cardboard-based physical media. Creative thinking was assessed using TTCT-inspired open-ended project design tasks aligned with four Guilford dimensions (fluency, flexibility, originality, elaboration), while student perceptions and qualitative learning experiences were examined through questionnaires and semi-structured interviews. Both groups demonstrated significant pre–post improvement; however, the plugged group achieved a higher N-Gain (0.489 vs. 0.419), both within the moderate category, with a statistically significant but practically moderate between-group difference ($p = .001$, Cohen's $d = 0.65$). Qualitative findings indicated that plugged coding supports iterative refinement and multimedia expression, whereas unplugged coding promoted collaboration and flexible thinking through physical co-construction. Perception data showed positive learning experiences in both groups, with perceived creativity support moderately correlated with N-Gain scores in both conditions. These findings suggest that both approaches effectively support creative thinking through distinct pedagogical mechanisms and provide practical guidance for designing inclusive and context-sensitive coding instruction in resource-constrained elementary settings. This study compared the two approaches as separate conditions; hybrid integration was not examined.

Keywords: Creative thinking skills; Coding instruction; Project-based learning, Scratch, Unplugged coding

How to Cite: Daulay, L. A., & Nurfitriani, R. (2026). Integrating Coding and Project-Based Learning to Foster Elementary Students' Creativity. *Prisma Sains: Jurnal Pengkajian Ilmu Dan Pembelajaran Matematika Dan IPA IKIP Mataram*, *14*(3), 1167–1186. https://doi.org/10.33394/j-ps.v14i3.19856

https://doi.org/10.33394/j-ps.v14i3.19856

INTRODUCTION

In today's rapidly evolving digital landscape, creative thinking has become one of the most essential competencies for students to navigate complex real-world challenges. Creativity, alongside critical thinking, collaboration, and communication, is widely recognized as a cornerstone of 21st-century education (Trilling & Fadel, 2009). Elementary school years represent a critical period for cultivating these skills, as creativity is malleable and can be systematically developed through exploration, inquiry, and problem-solving activities (Beghetto & Kaufman, 2014; Lucas & Spencer, 2017).

Coding education provides a promising pathway to foster creative thinking by developing computational thinking skills, including decomposition, pattern recognition, abstraction, and algorithmic design (Grover & Pea, 2013). In elementary settings, coding is typically delivered through two modalities: plugged coding, which uses digital tools such as Scratch, and unplugged coding, which employs physical media and kinesthetic activities without digital devices (Bell & Vahrenhold, 2018). Both approaches have shown potential in enhancing problem-solving and creative outcomes (Garcia-Marques et al., 2026; Lin et al., 2024; Sun et al., 2024).

Computational thinking has been widely recognized as a fundamental skill in K–12 education, encompassing problem decomposition, abstraction, and algorithmic thinking (Hsu et al., 2018; Polat & Yilmaz, 2022; Weintrop et al., 2016). Several studies have also emphasized that programming-based learning can support the development of higher-order thinking skills, including creativity and problem solving (Lye & Koh, 2014; Ng & Cui, 2021).

To optimize the impact of coding instruction, many scholars recommend embedding it within a student-centered framework such as Project-Based Learning (PjBL), which promotes higher-order thinking and student agency through sustained, authentic projects (Arif, 2024; Demir & Önal, 2021; Kokotsaki et al., 2016). Theoretically, plugged and unplugged coding stimulate different dimensions of creativity as conceptualized in Guilford (1967) framework (fluency, flexibility, originality, and elaboration). Plugged coding with Scratch's visual block interface facilitates rapid iteration, immediate feedback, and elaboration, thereby supporting fluency and originality in a low-floor, high-ceiling environment (de Ruiter & Bers, 2022; Resnick et al., 2009). Learning through programming environments such as Scratch enables students to construct meaningful digital artifacts and supports iterative learning processes, which are essential for creative development (Bers, 2020; Resnick et al., 2009). In contrast, unplugged coding encourages divergent and flexible thinking through kinesthetic and tactile engagement without technological constraints (Brackmann et al., 2017; Threekunprapa & Yasri, 2020).

Despite these theoretical advantages, empirical evidence comparing the effectiveness of plugged and unplugged coding implemented as standalone treatments within a PjBL framework remains limited, particularly with regard to their differential impact on Guilford's four dimensions of creative thinking. Most existing studies either examine the two approaches in isolation, focus primarily on computational thinking rather than creative thinking, or are conducted in well-resourced urban settings (Atmatzidou & Demetriadis, 2016; Kirçali & Özdener, 2023). There is a notable lack of quasi-experimental research in resource-constrained, non-metropolitan contexts, where digital infrastructure is limited.

Preliminary observations at both schools revealed limited coding infrastructure, including a computer-to-student ratio of 1:2, limited teacher training in coding instruction, and the absence of a systematic coding program. These conditions reflect the broader reality of many non-metropolitan Integrated Islamic elementary schools in Indonesia.

This study addresses the identified gap by employing a quasi-experimental design to compare the effectiveness of plugged coding (using Scratch) and unplugged coding (using physical media), each implemented independently in separate student groups, within a PjBL framework. The study was conducted at two Integrated Islamic Elementary Schools (SD IT) in Takengon, Central Aceh: SD IT An Najah and SD IT Al-Manar. The primary objective is to determine which instructional modality more effectively enhances each dimension of creative thinking fluency, flexibility, originality, and elaboration among fifth-grade students in this non-metropolitan setting.

METHOD

Research Design

This study employed a mixed-methods quasi-experimental design integrating quantitative and qualitative approaches to examine the effectiveness of plugged and unplugged coding instruction within a Project-Based Learning (PjBL) framework. A quasi-experimental design was selected because intact classes were used, making full individual randomization impractical. The quantitative strand was the dominant component, while the qualitative strand provided explanatory depth through a convergent, embedded integration strategy (Creswell & Clark, 2017).

To reduce institutional bias, a within-school random assignment procedure was applied. At each participating school, two existing Grade V classes were randomly assigned to either the experimental condition (PjBL with plugged coding using Scratch) or the control condition (PjBL with unplugged coding using physical media). This resulted in a Nonequivalent Control Group Design, as illustrated in Table 1.

Table 1. Research Design

Group	School	Pretest	Treatment	Posttest
Experimental (n = 50)	SD IT An-Najah & SD IT Al-Manar	O_1	X_1: PjBL + Plugged Coding (Scratch)	O_3
Control (n = 48)	SD IT An-Najah & SD IT Al-Manar	O_2	X_2: PjBL + Unplugged Coding (physical media/cardboard-based coding)	O_4

Both groups followed identical instructional objectives, time allocation, and PjBL phases, differing only in the coding modality used. This design strengthens internal validity by controlling for school-level differences. Baseline equivalence between groups was confirmed by an independent-samples t-test of pretest scores ($p > 0.05$), as reported in the Results section.

Participants

The study was conducted at two Integrated Islamic Elementary Schools (*Sekolah Dasar Islam Terpadu*/SD IT) in Takengon, Central Aceh Regency, namely SD IT An-Najah Takengon and SD IT Al-Manar Takengon during the even semester of the 2024/2025 academic year. Each school had exactly two Grade V classes, yielding a total population of 127 students across four classes.

A purposive sampling technique was applied to select 98 students based on three criteria: (1) active participation in classroom learning; (2) no prior experience with formal coding instruction; and (3) sufficient literacy and numeracy skills to complete assigned tasks. The final sample comprised 50 students in the experimental group and 48 in the control group, distributed across both schools. Students in both conditions worked in pairs to accommodate the 1:2 computer-to-student ratio.

Intervention

The intervention lasted six weeks, comprising 18 instructional meetings (2 × 35 minutes each) and separate pretest and posttest sessions. Both groups followed the same five-phase PjBL model adapted from (Kingston, 2018), differing only in the coding modality used.

Table 2. PjBL Implementation Phases, Activities, Tools/Media, and Expected Products

Phase	Meetings	Activities	Tools / Media	Expected Product
Entry Event & Questioning	1–2	Introduction to the driving question; brainstorming project ideas	Worksheets, discussion cards	Mind map of project ideas
Planning & Design	3–5	Designing project structure; assigning roles; creating storyboards or algorithm flowcharts	Storyboard templates; algorithm sequence cards	Project blueprint / algorithm flowchart
Prototyping	6–12	Building and testing initial project prototypes	Scratch 3.0 (plugged); cardboard grid board, arrow tokens, command cards (unplugged)	Interactive Scratch animation/game (plugged); physical cardboard-based coding board game (unplugged)

Phase	Meetings	Activities	Tools / Media	Expected Product
Revision & Refinement	13–16	Peer feedback and iterative revision of prototypes	Peer feedback rubric	Refined final project
Presentation & Reflection	17–18	Presenting final projects and guided reflection	Presentation checklist; reflection journal	Final project demonstration and reflection report

To control for teacher-related variability, the intervention in both conditions was delivered by the same instructor using standardized lesson plans across all classrooms**.**

Experimental Group: Plugged Coding with Scratch

Students in the experimental group used Scratch 3.0 (scratch.mit.edu), a free visual block-based programming platform. Project tasks involved creating interactive animated folktales, simple educational games, and simulations of Grade V science and social studies content. Core programming elements utilized included sprite customization, event-based triggers, loop and conditional logic blocks, and broadcast messaging. During the Prototyping and Revision phases, students were encouraged to iteratively test, modify, and elaborate their projects. Teacher facilitation focused on guiding students to incorporate multimedia elements: audio, animation, and interactive features and refine project functionality through continuous feedback cycles.

Control Group: Unplugged Coding with Physical Media (Cardboard-Based Coding)

Students in the control group engaged in coding activities using physical materials, including a grid-based movement board, directional arrow tokens, conditional command cards, and algorithm sequence strips. Project tasks required students to design and execute physical coding board games in which peers followed step-by-step programming instructions to navigate the grid and solve contextual problems. These activities emphasized kinesthetic learning, peer collaboration, and the externalization of algorithmic thinking through physical object manipulation. Iterative refinement occurred through peer discussion and physical rearrangement of instruction sequences without digital tools, making this approach fully viable in infrastructure-limited settings.

Research Instruments

Creative Thinking Skills Test

Creative thinking skills were assessed using an open-ended task instrument inspired by the Torrance Tests of Creative Thinking (Torrance, 1974) and structured according to Guilford (1967) four dimensions of creative thinking: fluency, flexibility, originality, and elaboration. The instrument was adapted to be age-appropriate for fifth-grade students and contextually relevant to the coding project activities conducted in this study. It consisted of four tasks, each designed to measure a dimension. Examples of the tasks are as follows:

- Fluency: “List as many uses as you can think of for a piece of cardboard in 5 minutes.”
- Flexibility: “Look at your list of ideas. How many different categories do they belong to? Name each category.”
- Originality: “Write one idea that you believe no other student would write. Explain why it is unique.”
- Elaboration: “Choose one idea and develop it into a detailed plan, including steps, materials needed, and possible outcomes.”

Table 3. Creative Thinking Skills Scoring Rubric (TTCT-Inspired)

Dimension	Score 1 (Low)	Score 2 (Developing)	Score 3 (Proficient)	Score 4 (Advanced)
Fluency	1–2 relevant responses	3–4 relevant responses	5–7 relevant responses	8+ relevant responses

Dimension	Score 1 (Low)	Score 2 (Developing)	Score 3 (Proficient)	Score 4 (Advanced)
Flexibility	1 category only	2 distinct categories	3 distinct categories	4+ distinct categories
Originality	Common idea; no reasoning	Slightly uncommon; minimal reasoning	Uncommon idea with brief justification	Highly unique idea with clear reasoning
Elaboration	No development	Minimal detail (1–2 steps)	Moderate detail (3–4 steps)	Comprehensive plan (5+ steps + materials + outcomes)

Originality scoring used a frequency-based approach: responses produced by fewer than 5% of the total sample (n = 98) were classified as original. Two independent raters scored the responses, and inter-rater reliability was measured using Cohen's Kappa (κ = 0.81). Discrepancies were resolved through deliberation until consensus was reached.

Classroom Observation Sheet

A structured observation instrument monitored PjBL implementation fidelity and student engagement across five indicators: (1) student task engagement, (2) collaborative interaction, (3) creative risk-taking, (4) use of tools and media, and (5) alignment with the designated PjBL phase. Each indicator was rated on a 4-point scale (1 = not observed to 4 = consistently observed) by a trained classroom teacher observer at each school during every instructional session. The fidelity score per session was calculated as the percentage of maximum possible points achieved. Inter-rater reliability between observers across both schools was assessed for a subsample of sessions (n = 6), yielding Cohen's κ = 0.78, indicating acceptable agreement.

Student Perception Questionnaire

A 15-item Likert-scale questionnaire (1 = Strongly Disagree to 5 = Strongly Agree) measured students' perceptions of the learning experience across four dimensions: enjoyment, perceived support for creativity, ease of tool use, and learning motivation. Sample items included: *"I feel more creative when working on this project"* and *"The tools used in this activity helped me generate new ideas."* The instrument demonstrated good internal consistency (Cronbach's α = 0.84).

Semi-Structured Interview Guide

A 10-question semi-structured interview protocol was used to explore students' subjective learning experiences, their perceived creative processes, and the challenges they encountered. Representative questions included: *"What part of the project felt most creative to you, and why?"* and *"How did using Scratch/the physical materials help you come up with new ideas?"* Interviews were conducted with 12 purposively selected students (6 from each group) representing three creativity levels (high, medium, and low) as determined by pretest scores. Each interview lasted approximately 15–20 minutes and was audio-recorded with participants' consent.

Instrument Validity and Reliability

Content validity for all instruments was established through expert judgment by two specialists in educational assessment and elementary education, each independently rating item relevance and clarity on a 4-point scale. Items with a Content Validity Index (CVI) below 0.78 were revised or removed prior to field testing. All instruments met the minimum reliability threshold of Cronbach's α ≥ 0.70.

Research Procedure

The research procedure was systematically organized to ensure consistency in implementation across both schools and to allow a clear comparison between the plugged and unplugged coding groups. Each stage was designed to support the validity of the intervention process, beginning with preparation and instrument validation, followed by baseline measurement, treatment implementation, post-intervention assessment, and qualitative follow-up through student interviews. The study was implemented in five sequential stages:

1. Preparation (two weeks prior): Development and expert validation of learning materials, lesson plans, and all instruments; coordination with school principals; structured instructor briefing to ensure consistent delivery across both schools.
2. Pretest (Week 0): Simultaneous administration of the creative thinking skills test at both schools to establish baseline equivalence between groups.
3. Treatment (Weeks 1–6): Concurrent implementation of PjBL with plugged coding (experimental classes) and PjBL with physical media (control classes) delivered in parallel at both SD IT An-Najah Takengon and SD IT Al-Manar Takengon, controlling for temporal confounds.
4. Posttest (Week 7): Simultaneous re-administration of the creative thinking skills test at both schools.
5. Interviews (Week 7, post-posttest): Semi-structured interviews with 12 selected students conducted on the same day as the posttest to minimize recall decay.

Data Analysis

Quantitative Analysis

Quantitative data were analyzed using SPSS Statistics Version 26 at $\alpha = 0.05$ through four stages: (1) descriptive statistics: mean, standard deviation, and Normalized Gain (N-Gain); (2) prerequisite tests: Shapiro–Wilk for normality, Levene's Test for homogeneity, and independent samples t-test for baseline equivalence; (3) within-group effectiveness: paired samples t-test for each group's pretest-to-posttest improvement; and (4) between-group comparison: independent samples t-test on N-Gain scores. Effect size was reported using Cohen's d, interpreted as small ($d < 0.2$), moderate ($0.5 \leq d < 0.8$), or large ($d \geq 0.8$), with 95% confidence intervals reported for all major comparisons.

Qualitative Analysis

Qualitative data from observations and interviews were analyzed using thematic analysis following Braun & Clarke (2006) six-phase framework: (1) familiarization, (2) initial code generation, (3) theme search, (4) theme review, (5) theme definition and naming, and (6) report production. Two researchers independently coded all data; inter-coder agreement was assessed, and discrepancies were resolved through deliberation until consensus was reached. Trustworthiness was established through methodological triangulation: combining interviews, observation, and quantitative data and member-checking, whereby key emerging themes were verified with two student informants after initial analysis to ensure accurate representation of their experiences (Lincoln & Guba, 1985).

Data Integration

Quantitative and qualitative findings were integrated using a convergent embedded strategy (Creswell & Clark, 2017). Qualitative findings were embedded within and interpreted alongside the dominant quantitative results. Integration was operationalized through a joint display in the discussion section, where N-Gain scores for each Guilford dimension were matched directly to thematic findings from interviews and observations. This approach enabled the identification of convergences and divergences between measured outcomes and students' reported creative processes.

Research Ethics

This study adhered to established ethical standards for research involving minors. Institutional permission was obtained from the principals of SD IT An-Najah Takengon and SD IT Al-Manar Takengon. Written informed consent was obtained from the parents or guardians of all 98 participating students, and verbal assent was obtained from the students before participation. Participation was voluntary, and students were free to withdraw at any time without penalty. All data were anonymized during analysis and reporting. Any visual materials included in this study were carefully selected and anonymized to protect student identity. Ethical approval for the study was granted by the Research Ethics Committee of IAIN Takengon before data collection.

RESULTS

Descriptive Statistics

Table 4 presents descriptive statistics of creative thinking scores for both groups before and after the intervention. The creative thinking test comprised four dimensions (fluency, flexibility, originality, elaboration), each scored on a 1–4 scale, yielding a maximum possible total score of 16 points. Raw scores were converted to a 100-point scale for reporting purposes.

Table 4. Complete Descriptive Statistics of Creative Thinking Scores

Statistic	Experimental (Plugged)		Gain	Control (Unplugged)		Gain
	Pretest	Posttest		Pretest	Posttest	
N	50	50	50	48	48	48
Mean	58.24	78.62	20.38	57.83	75.46	17.63
Median	58.50	79.00	20.00	58.00	75.50	17.50
Std. Deviation	8.45	7.92	4.21	8.72	8.35	4.08
Minimum	40	60	10	38	58	8
Maximum	76	96	32	75	92	30
N-Gain Score	—	—	0.489	—	—	0.419
N-Gain Category	—	—	*Moderate*	—	—	*Moderate*

Both groups showed comparable pretest means (Experimental: M = 58.24, SD = 8.45; Control: M = 57.83, SD = 8.72). After the six-week intervention, the experimental group showed a mean gain of 20.38 points (N-Gain = 0.489, Moderate category), while the control group showed a mean gain of 17.63 points (N-Gain = 0.419, Moderate category). Both groups, therefore, remained in the same improvement category, with an N-Gain difference of 0.070 between conditions.

Prerequisite Tests

Before hypothesis testing, the assumptions of normality and homogeneity of variance were examined, and pretest equivalence between groups was confirmed. Table 5 summarizes the results of all prerequisite tests.

Table 5. Results of Prerequisite Tests

Test	Group	Statistic	Df	p-value	Conclusion
Shapiro–Wilk (Pretest)	Experimental	0.981	50	0.612	Normal
	Control	0.977	48	0.503	Normal
Shapiro–Wilk (Posttest)	Experimental	0.975	50	0.374	Normal
	Control	0.979	48	0.541	Normal
Shapiro–Wilk (N-Gain)	Experimental	0.983	50	0.701	Normal
	Control	0.976	48	0.428	Normal
Levene's Test (Pretest)	Both	F = 0.184	—	0.669	Homogeneous
Levene's Test (N-Gain)	Both	F = 0.312	—	0.578	Homogeneous
Independent t-Test (Pretest)	Both	t = 0.235	96	0.815	No significant difference

All data were normally distributed (all Shapiro–Wilk $p > 0.05$) and variances were homogeneous (Levene's $F < 0.40$, $p > 0.05$). The pretest equivalence test confirmed no significant difference between groups at baseline ($t(96) = 0.235$, $p = 0.815$), supporting the validity of subsequent between-group comparisons.

Hypothesis Testing

Effectiveness of PjBL-Based Plugged Coding (H_1)

A paired-samples t-test was conducted to examine whether PjBL-based plugged coding instruction significantly improved students' creative thinking skills. Results showed a significant pre–post improvement, $t(49) = 34.087$, $p < 0.001$. The mean gain was 20.38 points (N-Gain = 0.489), with a large effect size, Cohen's d = 1.02, 95% CI for mean gain [19.17, 21.59]. These results support H_1: PjBL-based plugged coding instruction was effective in improving fifth-grade students' creative thinking skills.

Table 6. Paired Samples t-Test Results Experimental Group (Plugged Coding)

Comparison	**Mean Gain**	**SD**	**T**	**Df**	**p**	**Cohen's d**	**95% CI**
Pretest – Posttest	20.38	4.21	34.087	49	< 0.001	1.02	[19.17, 21.59]

Effectiveness of PjBL-Based Unplugged Coding (H_2)

A paired-samples t-test for the control group also showed a statistically significant pre–post improvement, $t(47) = 25.071$, $p < .001$. The mean gain was 17.63 points (N-Gain = 0.419), with a large effect size, Cohen's d = 0.89, 95% CI for mean gain [16.44, 18.82]. These results confirm that H_2 is supported: PjBL-based unplugged coding instruction was likewise effective in enhancing creative thinking skills.

Table 7. Paired Samples t-Test Results Control Group (Unplugged Coding)

Comparison	**Mean Gain**	**SD**	**T**	**Df**	**p**	**Cohen's d**	**95% CI**
Pretest - Posttest	17.63	4.08	25.071	47	< 0.001	0.89	[16.44, 18.82]

Comparative Effectiveness: Plugged vs. Unplugged Coding (H_3)

An independent samples t-test on N-Gain scores showed a statistically significant difference between groups, $t(96) = 3.284$, $p = .001$. The experimental group's N-Gain (M = 0.489) was significantly higher than the control group's (M = 0.419), with a moderate between-group effect size, Cohen's d = 0.65, 95% CI for the N-Gain difference [0.027, 0.113]. These results confirm that H_3 is supported: PjBL-based plugged coding produced significantly greater improvements in creative thinking than unplugged coding.

Table 8. Independent Samples t-Test on N-Gain Scores

	Experimental (Plugged)	**Control (Unplugged)**	**T**	**df**	**P**	**Cohen's d**	**95% CI (Difference)**
N-Gain Mean	0.489	0.419	3.284	96	0.001	0.65	[0.027, 0.113]

Creativity Improvement by Guilford's Dimensions

Table 9 presents pretest scores, posttest scores, gain scores, and the inter-group gain difference for each of the four Guilford (1967) creative thinking dimensions. All scores are reported on the original 1–4 scale per dimension.

Table 9. Creative Thinking Improvement by Dimension (Pretest, Posttest, Gain)

Dimension	**Experimental (Plugged)**			**Control (Unplugged)**			**Gain Difference**
	Pretest	**Posttest**	**Gain**	**Pretest**	**Posttest**	**Gain**	**(Exp − Con)**
Fluency	14.82	20.35	5.53	14.71	19.92	5.21	+0.32
Flexibility	14.36	19.88	5.52	14.29	19.24	4.95	+0.57
Originality	14.58	19.42	4.84	14.51	18.62	4.11	+0.73
Elaboration	14.48	18.97	4.49	14.32	17.68	3.68	+0.81
Total	58.24	78.62	20.38	57.83	75.46	17.63	+2.75

The experimental group showed greater gains than the control group across all four dimensions. The largest between-group difference was observed in elaboration (Δ = 0.81), followed by originality (Δ = 0.73), flexibility (Δ = 0.57), and fluency (Δ = 0.32). The smallest difference in fluency gains suggests both modalities were similarly effective in supporting the generation of multiple ideas. The pattern of larger differences in elaboration and originality is consistent with the total N-Gain advantage observed in the experimental group.

Student Perception Results

Students' perceptions of the learning experience were measured using a 10-item Likert-scale questionnaire (1–5 scale) administered at posttest. Table 10 presents mean scores and standard deviations for each perception dimension by group.

Table 10. Descriptive Statistics of Student Perception Questionnaire

Dimension	Experimental (Plugged)		Control (Unplugged)		Diff. (Exp – Con)
	M	**SD**	**M**	**SD**	**ΔM**
Enjoyment	4.32	0.61	4.15	0.68	+0.17
Creativity Support	4.41	0.57	4.18	0.64	+0.23
Ease of Use	4.20	0.65	4.05	0.71	+0.15
Motivation	4.38	0.59	4.22	0.63	+0.16
Overall Mean	4.33	0.61	4.15	0.67	+0.18

Both groups reported consistently positive perceptions of the learning experience, with overall mean scores of 4.33 (SD = 0.61) for the experimental group and 4.15 (SD = 0.67) for the control group. The experimental group scored slightly higher on all four dimensions, with the largest difference observed in perceived creativity support (ΔM = 0.23) and the smallest in ease of use (ΔM = 0.15).

Correlation Between Student Perceptions and Creative Thinking Improvement

To examine whether students' subjective perceptions of the learning experience were associated with objective gains in creative thinking, Pearson correlation analyses were conducted between overall perception scores and N-Gain scores within each group. Table 11 presents the correlation results

Table 11. Pearson Correlations Between Perception Scores and N-Gain Scores

Group	Perception Dimension	R	P	95% CI [LL, UL]	Interpretation
Experimental (n = 50)	Overall perception	0.423	.002	[0.168, 0.631]	Moderate, positive
	Creativity Support	0.461	.001	[0.212, 0.662]	Moderate, positive
	Enjoyment	0.387	.005	[0.126, 0.603]	Moderate, positive
	Motivation	0.401	.004	[0.143, 0.614]	Moderate, positive
Control (n = 48)	Overall perception	0.389	.006	[0.125, 0.607]	Moderate, positive
	Creativity Support	0.412	.004	[0.151, 0.626]	Moderate, positive
	Enjoyment	0.358	.013	[0.085, 0.585]	Small-to-moderate, positive
	Motivation	0.374	.009	[0.103, 0.597]	Moderate, positive

Note. r = Pearson correlation coefficient; LL = lower limit; UL = upper limit of 95% CI. Correlation strength: small = .10–.29, moderate = .30–.49, large ≥ .50 (Cohen, 1988). All correlations are statistically significant (p < .05). Perception scores = overall mean of the 10-item questionnaire (1–5 scale). N-Gain scores were used as the outcome variable in all correlations.

Significant positive correlations were found between overall perception scores and N-Gain scores in both the experimental group (r = 0.423, p = .002) and the control group (r = 0.389, p = .006). Within both groups, perceived creativity support showed the strongest association with N-Gain scores (Experimental: r = 0.461, p = .001; Control: r = 0.412, p = .004). These findings indicate that students who perceived the coding activities as more supportive of creative expression tended to demonstrate greater objective improvement in creative thinking skills, regardless of coding modality.

Qualitative Findings from Interviews and Observations

Qualitative data from semi-structured interviews with 12 purposively selected students (6 experimental, 6 control) and classroom observation field notes were analyzed using thematic analysis (Braun & Clarke, 2006). Three overarching themes emerged from the data. Participant codes are used throughout (S-E = experimental group student; S-K = control group student; numbers indicate participant identifier).

Theme 1: Iterative Exploration and Idea Development

This theme was identified in 10 of 12 interview transcripts (6 experimental, 4 control), making it the most prevalent theme overall. Students in the experimental group consistently emphasized the ability to test, revise, and immediately observe changes in their Scratch projects as a key feature of their learning experience. For example, one student stated: "In Scratch, I can try many ideas and change them quickly if they don't work" (S-E-01). Another noted: "Every time I add something, I run it to see if it looks right. Then I add more" (S-E-04). Classroom observations corroborated these accounts: experimental group students were observed making an average of 6–8 discrete iterative modifications per session during the prototyping phase (Meetings 6–12).

Although less prominent in the control group (4 of 8 interviewees), iteration was also present in unplugged activities, manifesting as physical rearrangement of instruction cards. One control group student described: "When our robot goes the wrong way, we change the cards and try again" (S-K-03). Observations noted that physical revision cycles were longer than digital ones, as students had to manually disassemble, reorder, and retest instruction sequences.

Theme 2: Collaboration and Peer Interaction

This theme appeared in 9 of 12 transcripts (4 experimental, 5 control), and was particularly prominent in the control group. Control group students highlighted direct peer discussion as central to their problem-solving process. One participant noted: "We can talk to our friends directly when organizing the steps" (S-K-02). Another described the collaborative nature of debugging physical sequences: "My partner and I argue sometimes about which step is wrong, but then we figure it out together" (S-K-06). Classroom observations showed that control group pairs engaged in verbal negotiation during algorithm design on average 4.2 episodes per session, compared to 2.8 in the experimental group, where on-screen work occasionally reduced the need for explicit verbalization.

Experimental group students also reported collaboration, though it was often structured around turn-taking at the computer: "We take turns. One types the blocks, the other watches and says what to fix" (S-E-07). This role differentiation was consistently observed across experimental pairs.

Theme 3: Observable Differences in Creative Expression

This theme was identified in 8 of 12 transcripts (5 experimental, 3 control) and was further supported by classroom observation field notes. Observed differences in how creativity was expressed across conditions were consistent with the quantitative patterns in Table 9. Experimental group students produced outputs with multiple visual and audio layers: painted backdrops, animated sprite sequences, sound effects, and multi-scene narratives. Observer field notes recorded that experimental group projects incorporated an average of 4.6 distinct multimedia elements by the end of the prototyping phase.

Control group students produced outputs demonstrating strong sequential logic and peer-verified algorithm correctness. Their projects required multiple conditional branches and were evaluated by peer testers who physically "ran" the coded sequences. Observer field notes recorded an average of 3.8 conditional branches per final board game design, reflecting strong flexibility in algorithmic thinking. One control group student described their work as: "We had to think of every possible way the player could move, so we needed lots of different instructions" (S-K-05).

Overall, the qualitative findings are consistent with and provide explanatory context for the quantitative results. The pattern of stronger elaboration and originality gains in the experimental group is supported by observations of richer multimedia outputs and more frequent iterative modifications. The comparable fluency gains across groups are consistent with the active idea generation observed in both conditions, whether digital or physical. Interpretation of these patterns in relation to theoretical frameworks is addressed in the Discussion section.

DISCUSSION

Effectiveness of PjBL-Based Plugged Coding Instruction

PjBL-based plugged coding using Scratch significantly improved students' creative thinking skills ($t = 34.087$; $p < 0.001$; N-Gain = 0.489; Cohen's $d = 1.02$), with the largest gains in fluency (5.53) and flexibility (5.52). These results are consistent with constructionist theory Papert (1980): students constructed shareable digital artifacts: animated stories, educational games, and science simulations that naturally engaged all four Guilford (1967) dimensions simultaneously. Scratch's low-floor, high-ceiling environment minimized technical barriers while enabling rapid experimentation, consistent with findings by Resnick et al. (2009), Sun et al. (2023), Yilmaz & İzmirli (2023). The PjBL framework amplified these effects by embedding coding within authentic, iterative project cycles rather than treating it as an isolated skill (Cui & Ng, 2021; Sri Utami, 2024).

These quantitative findings are corroborated by the qualitative and perception data collected in this study. Interview analysis (Section 3.6, Theme 1) revealed that students in the experimental group made more frequent iterative modifications during the prototyping phase, averaging 6–8 discrete revisions per session. Students described this iterative process explicitly: one participant noted, "In Scratch, I can try many ideas and change them quickly if they don't work" (S-E-01), while another described a continuous test-refine cycle: "Every time I add something, I run it to see if it looks right. Then I add more" (S-E-04). This pattern corresponds to the quantitative elaboration and originality gains (Table 9; Δ elaboration = 0.81; Δ originality = 0.73), suggesting that Scratch's immediate visual feedback mechanism functioned as a catalyst for iterative creative refinement, consistent with Resnick et al. (2009) creative learning spiral.

Perception data (Section 3.5) further reinforce this interpretation. Experimental group students reported the highest scores in perceived creativity support ($M = 4.41$, $SD = 0.57$) and motivation ($M = 4.38$, $SD = 0.59$) among all perception dimensions. Critically, perceived creativity support showed the strongest correlation with N-Gain scores in the experimental group ($r = 0.461$, $p = .001$), indicating that students who felt more supported in creative expression also achieved greater objective improvement. The convergence of these data sources strengthens the evidence for the observed effect.

However, the interpretation of these results must acknowledge a potential novelty effect: as students at both schools had no prior exposure to Scratch, elevated engagement and originality scores may partly reflect the excitement of encountering a new digital environment rather than the platform's inherent instructional superiority alone. Educational technology research consistently documents this phenomenon, whereby new tools temporarily boost performance due to perceived attractiveness (Clark, 1983; Mayer, 2014). Perception data provide indirect support for this concern: experimental group students reported higher enjoyment scores ($M = 4.32$ vs. $M = 4.15$), and while enjoyment correlated positively with N-Gain ($r = 0.387$, $p = .005$), the moderate magnitude of this correlation suggests that novelty-driven engagement was a contributing but not dominant factor. Future studies should include delayed posttest measurements to determine whether gains persist beyond the initial exposure period.

This finding is consistent with prior research indicating that programming environments support iterative development and creative expression through continuous testing and modification of ideas (Grover & Pea, 2013; Hsu et al., 2018).

Effectiveness of PjBL-Based Unplugged Coding Instruction

PjBL-based unplugged coding also produced significant improvements ($t = 25.071$; $p < 0.001$; N-Gain = 0.419; Cohen's $d = 0.89$), demonstrating that device-free instruction can meaningfully develop creative thinking in resource-constrained settings. The effectiveness is explained through Piaget's concrete operational theory: manipulating physical grid boards, arrow tokens, and command cards allowed students to externalize abstract algorithmic concepts through direct sensorimotor engagement, developmentally appropriate for Grade V students (Dağ et al., 2023; Threekunprapa & Yasri, 2020). Lower absolute gains compared to the plugged group likely reflect the higher physical effort required per revision cycle, which constrained iterative elaboration.

Teacher scaffolding played a critical role in the unplugged condition. The physical nature of cardboard-based activities requires active facilitation, the management of materials, support for sequencing, and sustained engagement, making learning outcomes more sensitive to facilitation quality than in the digital environment. Vygotsky (1978) Zone of Proximal Development framework highlights this dynamic: effective teacher scaffolding bridges the gap between students' current ability and their potential creative output, a process that is more explicitly visible and more variable in physical coding activities. Future studies should implement structured inter-rater fidelity protocols to more rigorously quantify and control teacher influence across conditions.

Qualitative findings (Section 3.6, Theme 2) directly support this explanation. Control group students consistently emphasized collaboration and direct peer interaction as central to their learning experience. One participant described: "We can talk to our friends directly when organizing the steps" (S-K-02), while another highlighted the productive tension of peer debugging: "My partner and I argue sometimes about which step is wrong, but then we figure it out together" (S-K-06). Classroom observations corroborated this pattern: control group pairs engaged in verbal negotiation an average of 4.2 episodes per session, compared to 2.8 in the experimental group. This suggests that creativity in unplugged settings is primarily facilitated through social co-construction rather than individual rapid iteration, a distinction consistent with Vygotsky's emphasis on the social origins of higher mental functions.

Perception data provide additional supporting evidence. Control group students reported positive perceptions across all dimensions (overall $M = 4.15$, $SD = 0.67$), with motivation receiving the highest mean score ($M = 4.22$). The moderate positive correlation between perceived creativity support and N-Gain scores in the control group ($r = 0.412$, $p = .004$) indicates that students who perceived the physical activities as creatively supportive also achieved greater objective gains, mirroring the pattern observed in the experimental group. The comparable correlation magnitudes across groups ($r = 0.461$ vs. $r = 0.412$) suggest that the perception–outcome association is robust across coding modalities, not specific to the digital environment.

Comparative Effectiveness and Practical Implications

Plugged coding demonstrated statistically greater effectiveness than unplugged coding ($t = 3.284$, $p = 0.001$, Cohen’s $d = 0.65$), with the most pronounced advantages in originality ($\Delta = 0.73$) and elaboration ($\Delta = 0.81$). These differences are attributable to Scratch's multimedia affordances and immediate visual feedback, which are known to support iterative creative processes (Grover & Pea, 2013; Resnick et al., 2009). As discussed in Section 4.1, a novelty effect may have contributed to these advantages and should be considered when interpreting the between-group difference (Clark, 1983; Mayer, 2014).

Despite this statistically significant advantage, the practical difference remained moderate (Cohen's d = 0.65; N-Gain difference = 0.070). Both approaches achieved the same N-Gain category (Moderate), and fluency gains were relatively comparable (Δ = 0.32). Perception and correlation findings (Section 3.5) provide further convergent evidence, both groups reported similarly positive learning experiences (overall perception M = 4.33 vs. M = 4.15), and the strength of the perception–N-Gain correlation was comparable across conditions (r = 0.423 vs. r = 0.389). This convergence suggests that the subjective quality of the learning experience, rather than modality alone, plays a meaningful role in the development of creative thinking. Educators implementing either approach should therefore attend to how students perceive the creative affordances of their activities, rather than only to the chosen modality.

Qualitative data (Section 3.6, Theme 3) further contextualize the quantitative gap. Experimental group students produced richer multimedia outputs, layered animations, interactive game mechanics, and audio integration, averaging 4.6 distinct multimedia elements per final project. These characteristics correspond directly to the quantitative advantages observed in elaboration and originality. Control group students, by contrast, produced algorithmically complex board games averaging 3.8 conditional branches per design, reflecting strong flexibility in sequential logic. This artifact-level evidence, while descriptive rather than formally scored, is consistent with the dimension-level quantitative patterns in Table 9 and supports the interpretation that the two modalities engage different aspects of creative thinking, a distinction not captured by total N-Gain scores alone.

These findings indicate that unplugged coding remains a viable and effective approach, particularly in contexts with limited device access, a concern of growing importance in equitable computing education (Margolis & Nao, 2008). Previous studies have highlighted that both plugged and unplugged approaches can contribute to the development of computational thinking, though through different mechanisms of engagement (Lye & Koh, 2014; Weintrop et al., 2016). The results of this study extend these findings by demonstrating differential effects at the level of specific creativity dimensions, providing more granular guidance for instructional design decisions.

The role of teacher scaffolding, theoretically framed through Vygotsky (1978) Zone of Proximal Development, is particularly relevant when comparing conditions. Both groups demonstrated high implementation fidelity (84.4% and 83.4%), yet the qualitative data suggest that scaffolding functioned differently across modalities: in the plugged condition, scaffolding primarily prompted reflection and iteration; in the unplugged condition, it additionally managed physical materials and sustained collaborative engagement. Teacher scaffolding has been widely recognized as a key factor influencing student creativity and problem-solving across contexts (Sawyer & Henriksen, 2023; Vygotsky, 1978).

These findings may extend beyond the specific context of Integrated Islamic elementary schools in Central Aceh to other resource-constrained, non-metropolitan, and developing-region settings with comparable infrastructure constraints. The consistent effectiveness of both modalities suggests that PjBL-based coding instruction is a robust pedagogical approach adaptable across varying levels of technological access.

Cross-Site Consistency and Artifact Evidence

N-Gain improvements were consistent across both participating schools. Experimental classes at SD IT An-Najah and SD IT Al-Manar showed comparable gains, as did control classes at both sites. This cross-site consistency strengthens confidence that the observed treatment effects are primarily associated with the coding modality rather than with site-specific factors such as teacher personality, classroom culture, or local resources (Fixsen et al., 2005; Kokotsaki et al., 2016).

Student-produced artifacts provided additional performance-based evidence of creative development. Experimental group Scratch projects featured layered animations, interactive game mechanics, and audio integration, reflecting the observed quantitative gains in elaboration and originality. The control group's cardboard-based board games incorporated progressively more complex conditional instruction sequences, consistent with the fluency and flexibility gains observed in their quantitative data. These artifact patterns align with qualitative Theme 3 findings (Section 3.6.3), which documented an average of 4.6 multimedia elements in experimental projects, compared with 3.8 conditional branches in control group designs.

Perception data add a further layer of convergent evidence: students in the experimental group who reported higher creativity support scores also tended to produce richer multimedia outputs, consistent with the perception–N-Gain correlation (r = 0.461). Similarly, control group students' positive motivation scores (M = 4.22) may reflect the intrinsic satisfaction of successfully debugging a physical instruction sequence, a process associated with flexibility gains. Artifact-based assessment is widely recognized as an important complement to test-based evaluation in creative learning contexts (Brennan & Resnick, 2012; Papert, 1980).

Figures 1–3 present visual evidence from student artifacts and classroom activities.

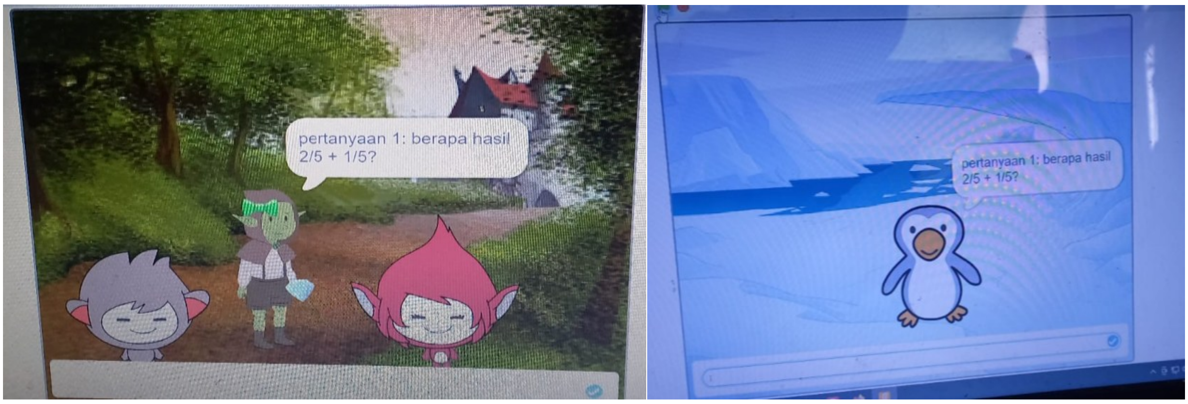

Figure 1. Example of a student-developed Scratch project featuring animated characters and interactive elements

Figure 2. Students working with Scratch block-based programming, demonstrating the use of loop structures

Figure 3. Students collaboratively engaging in unplugged coding activities using physical materials

However, formal systematic artifact analysis was not conducted in this study; therefore, interpretations based on student work remain descriptive and illustrative rather than inferential. Future research should incorporate structured artifact evaluation using validated rubrics to provide stronger performance-based evidence of the development of creative thinking (Brennan & Resnick, 2012).

In addition to its empirical findings, this study contributes a structured, replicable instructional model integrating PjBL with both plugged-in and unplugged coding modalities, alongside a contextually adapted creative thinking rubric based on Guilford (1967) dimensions. This methodological framework may serve as a practical reference for educators and researchers implementing coding-based interventions in elementary education across varying levels of technological access.

Mixed-Methods Synthesis: Convergence of Quantitative, Qualitative, and Perception Evidence

This study employed a convergent embedded mixed-methods design (Creswell & Clark, 2017), in which qualitative interview data and student perception scores were collected concurrently with the quantitative outcome data and used to explain and contextualize quantitative patterns. The overall pattern of findings across all three data strands is convergent, strengthening confidence in the study's conclusions.

At the level of between-group differences, all three data strands point in the same direction: the experimental group demonstrated greater creative thinking improvement quantitatively (N-Gain 0.489 vs. 0.419; Cohen's $d = 0.65$), reported more positive perceptions of creativity support and motivation, and produced qualitatively richer artifacts with more frequent iterative modifications. The convergence of these three sources provides stronger evidence for the advantage of plugged coding in this context than any single data strand could provide alone.

At the within-group level, the two modalities appear to engage creative thinking distinct processes that emphasize different aspects of creative thinking. For the plugged condition, the dominant mechanism appears to be individual iterative development enabled by immediate digital feedback, a pattern supported by interview Theme 1, elaboration/originality gain advantages, and the perception–N-Gain correlation for creativity support ($r = 0.461$). For the unplugged condition, the dominant mechanism appears to be collaborative co-construction through verbal negotiation, a pattern supported by interview Theme 2, observation data (4.2 verbal negotiation episodes per session), and the comparable motivation perception scores. These distinct mechanisms explain why fluency gains were similar across conditions (idea generation benefits from both individual iteration and collaborative brainstorming), while elaboration and originality gaps were larger (these dimensions benefit specifically from the rapid feedback and multimedia affordances of the digital environment).

One area of divergence across data strands warrants further consideration. While quantitative data show a statistically significant between-group difference, perception data suggest that both groups experienced similarly positive learning environments ($\Delta M = 0.18$ on overall perception), and qualitative data indicate that control group students derived genuine creative satisfaction from physical activities. This divergence is theoretically informative: it suggests that objective creative performance and subjective experience of creativity do not map perfectly onto each other, and that unplugged activities can be highly engaging and motivating even when they yield slightly lower test-based gains. This finding has direct practical implications: in resource-constrained settings, unplugged coding is not merely a compromise; it represents a pedagogically legitimate and motivationally effective approach in its own right.

However, the effectiveness of combining these modalities within a hybrid instructional design was not examined in this study. Overall, these findings provide convergent evidence that PjBL-based coding instruction is effective and adaptable across varying levels of technological access.

Limitations

Five key limitations should be considered when interpreting these findings. First, a novelty effect may have temporarily elevated engagement and performance in the experimental group, as students had no prior exposure to Scratch. Although perception–N-Gain correlation data suggest novelty was a contributing rather than dominant factor ($r = 0.387$ for enjoyment vs. $r = 0.461$ for creativity support), delayed posttest measurements are recommended in future studies to assess whether gains persist beyond the initial exposure period.

Second, the 1:2 device ratio at both schools required pair work in the plugged condition, potentially moderating individual creative output and introducing a social dynamic not present in a one-to-one computing setting (Margolis & Nao, 2008). The extent to which pair work constrained or enhanced individual creative performance could not be disaggregated in the current design.

Third, researcher influence cannot be fully excluded, as the lead researcher delivered all sessions across both conditions and schools. Although implementation fidelity was monitored (84.4% and 83.4%) and a standardization briefing was conducted, independent teacher delivery in a future study would strengthen causal claims (Campbell & Stanley, 1965).

Fourth, originality scoring was not conducted under fully blinded conditions despite strong inter-rater reliability ($\kappa = 0.81$), introducing a minor risk of assessment bias, particularly for the originality dimension, where frequency-based scoring required reference to the full sample distribution.

Fifth, the qualitative strand, while systematically conducted ($\kappa = 0.78$, member-checking confirmed), was based on a purposive sample of 12 students and should be interpreted cautiously. High- and low-gain student perspectives were represented, but other sources of variation (gender, learning style, prior academic performance) were not systematically controlled for. All findings should be interpreted within the specific non-metropolitan SD IT context of Central Aceh, and caution is warranted when generalizing to urban, better-resourced, or culturally different settings.

CONCLUSION

This study examined the comparative effectiveness of plugged and unplugged coding instruction in a Project-Based Learning (PjBL) framework. The focus was on creative thinking skills in fifth-grade students at non-metropolitan Integrated Islamic elementary schools in Central Aceh, Indonesia. The study used quantitative outcomes, qualitative interviews, student perception scores, and classroom artifacts. Three main conclusions emerge.

First, both instructional modalities improved creative thinking across all four Guilford (1967) dimensions: fluency, flexibility, originality, and elaboration. Large within-group effect sizes were seen (Cohen's $d = 1.02$ and 0.89). These findings indicate that PjBL-based coding instruction, either digital or non-digital, is viable for elementary students. This is especially true in non-metropolitan schools, where resources are fewer.

Second, plugged coding was more effective than unplugged coding, especially in elaboration and originality. The between-group effect was moderate (Cohen's d = 0.65). Scratch's multimedia features and immediate feedback may explain this difference. Interviews showed more frequent iterative modifications. They also showed higher perceived creativity-support scores in the experimental group. A novelty effect may have influenced this advantage. Future studies should use delayed posttests to investigate further.

Third, the two modalities appear to engage creative thinking through different mechanisms. Plugged coding emphasizes individual iterative development. Unplugged coding, in contrast, emphasizes collaborative co-construction. Both groups reported similarly positive learning experiences. Strong correlations were also observed between perception and performance. These findings indicate unplugged coding is a valid pedagogical choice, especially where digital infrastructure is limited.

These findings have practical implications for educators and curriculum designers. Unplugged coding is accessible and effective, even in environments with limited technology. Plugged coding can be introduced gradually to support more advanced creative expression. Effective use of either approach requires structured teacher training in coding and PjBL facilitation. Extra attention to scaffolding is needed in unplugged conditions.

It is important to note that this study compared the two modalities as separate conditions rather than a hybrid model. The finding that the two approaches operate through different mechanisms suggests that a sequenced unplugged-to-plugged design may represent a promising direction for future research; however, this was not examined in the present study.

RECOMMENDATION

Based on the findings, elementary school teachers are encouraged to integrate coding activities into Project-Based Learning to foster students' creative thinking skills. Plugged coding using Scratch can be prioritized when digital facilities are available because it supports iterative refinement, multimedia expression, originality, and elaboration. However, unplugged coding using physical media should also be considered a viable alternative in resource-constrained schools, as it promotes collaboration, flexible thinking, and active engagement without requiring digital devices. Schools are also encouraged to provide teacher training on coding-based PjBL implementation, including project design, scaffolding strategies, and creativity assessment. Future studies should examine hybrid or sequenced unplugged-to-plugged coding models, involve broader school contexts, and include delayed posttests to determine the sustainability of students' creative thinking gains.

ACKNOWLEDGMENTS

The researchers would like to express their gratitude to all parties involved in this study. The researchers also extend their appreciation to the State Islamic Institute (IAIN) Takengon for its support in providing research funding and publication assistance for this study.

FUNDING INFORMATION

This study received research funding and publication assistance from the State Islamic Institute (IAIN) Takengon.

AUTHOR CONTRIBUTIONS STATEMENT

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Leni Agustina Daulay	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
Rahmah Nurfitriani	✓	✓	✓	✓			✓	✓		✓	✓			

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

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

REFERENCES

Arif, M.K. (2024). Enhancing project-based learning in STEM education with integrated technology and coding. *Journal of Intelligent Systems and Information Technology, 1*(1), 2024. https://jurnalapik.id/index.php/jisit

Atmatzidou, S., & Demetriadis, S. (2016). Advancing students' computational thinking skills through educational robotics: A study on age and gender relevant differences. *Robotics and Autonomous Systems, 75*, 661–670. https://doi.org/10.1016/j.robot.2015.10.008

Beghetto, R. A., & Kaufman, J. C. (2014). Classroom contexts for creativity. *High Ability Studies, 25*(1), 53–69. https://doi.org/10.1080/13598139.2014.905247

Bell, T. C., & Vahrenhold, J. (2018). CS unplugged: How is it used, and does it work? *Adventures Between Lower Bounds and Higher Altitudes*. https://api.semanticscholar.org/CorpusID:52166808

Bers, M. (2020). *Coding as a playground*. Routledge. https://doi.org/10.4324/9781003022602

Brackmann, C. P., Moreno-León, J., Román-González, M., Casali, A., Robles, G., & Barone, D. (2017). Development of computational thinking skills through unplugged activities in primary school. *ACM International Conference Proceeding Series*, 65–72. https://doi.org/10.1145/3137065.3137069

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. *AERA*, 1–12.

Campbell, D. T., & Stanley, J. C. (1965). Handbook of research on teaching. *British Journal of Educational Studies, 14*(1). https://doi.org/10.2307/3119055

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*(4), 445–459. https://doi.org/10.3102/00346543053004445

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

Cui, Z., & Ng, O. L. (2021). The interplay between mathematical and computational thinking in primary school students' mathematical problem-solving within a programming environment. *Journal of Educational Computing Research, 59*(5), 988–1012. https://doi.org/10.1177/0735633120979930

Dağ, F., Şumuer, E., & Durdu, L. (2023). The effect of an unplugged coding course on primary school students' improvement in their computational thinking skills. *Journal of Computer Assisted Learning, 39*(6), 1902–1918. https://doi.org/10.1111/jcal.12850

de Ruiter, L. E., & Bers, M. U. (2022). The coding stages assessment: Development and validation of an instrument for assessing young children's proficiency in the ScratchJr programming language. *Computer Science Education, 32*(4), 388–417. https://doi.org/10.1080/08993408.2021.1956216

Demir, C. G., & Önal, N. (2021). The effect of technology-assisted and project-based learning approaches on students' attitudes towards mathematics and their academic achievement. *Education and Information Technologies, 26*(3), 3375–3397. https://doi.org/10.1007/s10639-020-10398-8

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. University of South Florida.

Garcia-Marques, M. E., Pérez-Suay, A., & García-Bayona, I. (2026). Plugged or unplugged? A comparative study of computational thinking development in early childhood. *Education Sciences, 16*(2), 1–17. https://doi.org/10.3390/educsci16020333

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

Guilford, J. P. (1967). Creativity: Yesterday, today and tomorrow. *The Journal of Creative Behavior, 1*(1), 3–14. https://doi.org/10.1002/j.2162-6057.1967.tb00002.x

Hsu, T.-C., Chang, S.-C., & Hung, Y.-T. (2018). How to learn and how to teach computational thinking: Suggestions based on a review of the literature. *Computers & Education, 126*, 296–310. https://doi.org/10.1016/j.compedu.2018.07.004

Kingston, S. (2018). Project based learning and student achievement: What does the research tell us? *Buck Institute for Education, 1*(1), 1–11. http://bie.org/x9JN

Kirçali, A. Ç., & Özdener, N. (2023). A comparison of plugged and unplugged tools in teaching algorithms at the K-12 level for computational thinking skills. *Technology, Knowledge and Learning, 28*(4), 1485–1513. https://doi.org/10.1007/s10758-021-09585-4

Kokotsaki, D., Menzies, V., & Wiggins, A. (2016). Project-based learning: A review of the literature. *Improving Schools, 19*(3), 267–277. https://doi.org/10.1177/1365480216659733

Lin, Y., Liao, H., Weng, S., & Dong, W. (2024). Comparing the effects of plugged-in and unplugged activities on computational thinking development in young children. *Education and Information Technologies, 29*(8), 9541–9574. https://doi.org/10.1007/s10639-023-12181-x

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.

Lucas, B., & Spencer, E. (2017). *Teaching creative thinking: Developing learners who generate ideas and can think critically*. https://api.semanticscholar.org/CorpusID:210362402

Lye, S. Y., & Koh, J. H. L. (2014). Review on teaching and learning of computational thinking through programming: What is next for K-12? *Computers in Human Behavior, 41*, 51–61. https://doi.org/10.1016/j.chb.2014.09.012

Margolis, J., & Nao, K. (2008). *Stuck in the shallow end*. The MIT Press.

Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning* (2nd ed.). Cambridge University Press.

Ng, O. L., & Cui, Z. (2021). Examining primary students' mathematical problem-solving in a programming context: Towards computationally enhanced mathematics education. *ZDM: Mathematics Education, 53*(4), 847–860. https://doi.org/10.1007/s11858-020-01200-7

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books.

Polat, E., & Yilmaz, R. M. (2022). Unplugged versus plugged-in: Examining basic programming achievement and computational thinking of 6th-grade students. *Education and Information Technologies, 27*(7), 9145–9179.

Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., & Kafai, Y. (2009). Scratch: Programming for all. *Communications of the ACM, 52*(11), 60–67. https://doi.org/10.1145/1592761.1592779

Sawyer, R. K., & Henriksen, D. (2023). *Explaining creativity: The science of human innovation*. Oxford University Press. https://doi.org/10.1093/oso/9780197747537.001.0001

Sri Utami. (2024). A critical appraisal: Elementary school coding education effectiveness with Scratch in alignment with cognitive development, analyzed through Piagetian lens. *Journal of Electrical Systems, 20*(5s), 1180–1187. https://doi.org/10.52783/jes.2432

Sun, L., Guo, Z., & Zhou, D. (2023). Measuring development of young students' coding ability through a graphical teaching intervention: Further explanation of the effect of coding experience and coding interest. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2023.2190356

Sun, L., Liu, J., & Liu, Y. (2024). Comparative experiment of the effects of unplugged and plugged-in programming on computational thinking in primary school students: A perspective of multiple influential factors. *Thinking Skills and Creativity, 52*, Article 101542. https://doi.org/10.1016/j.tsc.2024.101542

Threekunprapa, A., & Yasri, P. (2020). Patterns of computational thinking development while solving unplugged coding activities coupled with the 3S approach for self-directed learning. *European Journal of Educational Research, 9*(3), 1025–1045. https://doi.org/10.12973/eu-jer.9.3.1025

Torrance, E. P. (1974). *Torrance tests of creative thinking*. Educational and Psychological Measurement.

Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. John Wiley & Sons.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology, 25*(1), 127–147. https://doi.org/10.1007/s10956-015-9581-5

Yilmaz, T., & İzmirli, S. (2023). Effect of unplugged and plugged coding activities on secondary school students' computational thinking skills. *Journal of Educational Technology and Online Learning, 6*(4), 1180–1193. https://doi.org/10.31681/jetol.1375335