MULTIDISCIPLINARY SCIENCE JOURNAL

RESEARCH ARTICLE
Published Online: May 31, 2026
https://doi.org/10.31893/multiscience.2026740

ARICT: An innovative learning model to enhance students' computational thinking using APOS–RME and digital learning integration

Trisnawati[a] | Sugeng Sutiarso[a] | Rangga Firdaus[a] | Tina Yunarti[a]

[a]Doctoral Program in Education, Faculty of Teacher Training and Education, University of Lampung, Bandarlampung, Indonesia.

Abstract Computational thinking (CT) is increasingly positioned as a transferable competence required for complex problem solving in mathematics and other STEM domains, yet CT instruction is often implemented through fragmented activities rather than coherent, theory-driven learning models. This study developed and evaluated ARICT, an innovative learning model that integrates APOS-informed cognitive construction, Realistic Mathematics Education (RME) contextualization, and digital learning supports to strengthen students' CT while maintaining disciplinary understanding. The research adopted a research and development approach aligned with the Borg and Gall model, implemented within a mixed methods paradigm. After needs analysis and evidence mapping, the ARICT package was produced, validated by experts, refined through a pilot study, and tested through a nonequivalent control group pretest–posttest field trial. The main study involved four intact Linear Algebra classes focusing on vectors (N = 80), comprising an experimental group (N = 41) receiving ARICT and a control group (N = 39) receiving conventional instruction. Baseline equivalence was supported by nonsignificant pretest differences ($p > .05$). Effectiveness was evaluated using ANCOVA with pretest scores as covariates. The results showed that ARICT significantly improved CT compared with conventional instruction, $F(1, 77) = 98.67$, $p < .001$, with higher adjusted posttest means for the experimental group ($M_{adj} = 78.52$) than the control group ($M_{adj} = 61.34$) and a large effect (partial $\eta^2 = .56$). ARICT also significantly improved vector conceptual understanding, $F(1, 77) = 45.21$, $p < .001$, with adjusted means of 82.15 (experimental) and 70.98 (control) and a large effect (partial $\eta^2 = .37$). Fidelity observations indicated high adherence to intended instructional conditions in both groups, supporting internal validity. These findings suggest that ARICT provides a coherent, scalable instructional model for embedding CT within mathematics learning, yielding strong CT gains without sacrificing conceptual mastery.

Keywords: computational thinking, digital learning, vector concepts, instructional model, higher education

1. Introduction

Contemporary education increasingly emphasizes transferable competencies that enable learners to interpret complex information, construct systematic solutions, and justify decisions across contexts. Within this landscape, computational thinking (CT) has become a widely discussed competency because it supports structured reasoning, model-based problem solving, and the design of solution procedures that can be tested and refined. Research in mathematics education shows strong momentum in integrating CT into mathematics learning, supported by the argument that mathematics offers natural entry points for CT through problem solving, representation, and reasoning practices (Kallia et al., 2021).

In applied educational research, CT is commonly operationalized as a set of teachable and assessable components. A validated measurement approach conceptualizes CT through core dimensions such as abstraction, decomposition, algorithmic thinking, evaluation, and generalization, enabling CT to be treated as a multidimensional construct rather than an abstract slogan (Tsai et al., 2021). Recent synthesis work also indicates that CT assessment in higher education remains methodologically diverse and that many studies combine multiple methods to strengthen measurement credibility and interpretability (X. Zhang et al., 2024). These developments imply that CT should be approached as a planned learning outcome that requires explicit instructional design and coherent assessment alignment.

Despite increasing attention, implementing CT in mathematics instruction is not straightforward. A literature-informed Delphi study highlights that CT in mathematics education intersects with problem solving, cognitive processes, and instructional transposition and that practical integration requires careful coordination between computational ideas, mathematical meaning, and classroom realities (Serpe & Frassia, 2020). A systematic review focusing on primary mathematics reports that CT-related classroom activities often cluster into skill-oriented approaches and process-oriented approaches, whereas comprehensive integration that systematically links CT practices to mathematical learning trajectories remains limited (Nordby

Received: January 21, 2026 | Accepted: March 11, 2026

et al., 2022). Although this synthesis targets primary contexts, it indicates a broader pattern that is also relevant for higher education and teacher education: CT is frequently introduced through isolated tasks rather than through a coherent pedagogical system.

Technology adoption alone does not guarantee CT growth. Recent meta-level syntheses report that CT interventions vary widely in design and effectiveness and that stronger effects are typically associated with structured pedagogical strategies rather than the mere presence of digital tools (Zhanga et al., 2025). A systematic review of CT for the digital age similarly emphasized that the field has expanded rapidly, yet the effectiveness of CT development depends on the alignment among learning tools, instructional strategies, and assessment practices, including explicit design choices that support learning progression (Rao & Bhagat, 2024). This motivates a shift from tool-focused innovation toward model-driven innovation, where instructional phases, task sequences, and scaffolding mechanisms are specified and empirically tested.

The present study is grounded in teacher education, where CT development has direct implications for future classroom practice. A systematic review of preservice teachers' CT indicates that structured training can improve CT, emphasizing that teacher education needs planned and coherent interventions to support CzT development and its transfer to instructional design (Dong et al., 2024). In the local context of this study, preliminary diagnostic evidence indicates that prospective mathematics teacher students' CT remains only moderate across components. Beyond general computational thinking difficulties, the challenge becomes particularly visible in the learning of vectors in linear algebra . Vector topics require students to coordinate multiple representations including algebraic notation, geometric interpretation, and procedural operations such as vector addition, scalar multiplication, and direction analysis. Research in mathematics education indicates that students frequently struggle to translate conceptual reasoning about vectors into systematic procedures, often relying on memorized formulas without understanding the underlying structure of vector operations. This difficulty becomes more pronounced when solving contextual or multistep vector problems that require decomposition of spatial relationships and algorithmic reasoning. Previous studies also report that conventional instruction in linear algebra tends to emphasize symbolic manipulation rather than structured reasoning processes, which limits students' opportunities to practice computational thinking components such as decomposition, abstraction, and algorithm design. As a result, vector learning often remains procedural and fragmented, preventing students from developing coherent strategies for analyzing and solving complex vector problems .

These characteristics suggest that vectors constitute a mathematically appropriate yet cognitively demanding domain for examining computational thinking development. Without structured pedagogical support, students may fail to connect conceptual understanding with systematic solution procedures, particularly in algorithm design and evaluation. Therefore, instructional models that explicitly integrate computational thinking processes with conceptual mathematical learning may be necessary to address these persistent difficulties in vector learning. In an initial assessment of 73 students via a 1–4 scoring scale, the mean scores were 2.4 for decomposition, 2.7 for pattern recognition, 2.5 for abstraction, 2.3 for algorithm design, and 2.6 for evaluation. This profile suggests that algorithm design is a particularly critical weakness because it represents students' ability to translate reasoning into systematic procedures that can be executed and checked.

The ARICT learning model was developed as an integrated pedagogical framework that synthesizes APOS theory, realistic mathematics education (RME), and digital learning to address the need for coherence in computational thinking-oriented mathematics instruction. This integration is grounded in the recognition that each approach offers substantial strengths but also has limitations when applied in isolation. APOS theory provides a robust cognitive explanation of how mathematical understanding develops through actions, processes, objects, and schemas; however, it offers limited guidance for consistent instructional enactment in classroom contexts (Herawaty et al., 2020; Palobo et al., 2025). In parallel, RME emphasizes that mathematical learning should begin from meaningful and experientially accessible contexts, enabling learners to construct formal concepts through progressive mathematization rather than immediate abstraction (Fredriksen, 2020; Gravemeijer, 2020). However, RME-based implementations are often not systematically aligned with explicit computational thinking indicators.

Moreover, digital learning environments offer rich affordances for exploration, visualization, interaction, and reflection. Nevertheless, research consistently shows that technology use alone does not guarantee meaningful learning or computational thinking development when it is not embedded within a coherent pedagogical model (Bozkurt et al., 2022; Engelbrecht & Borba, 2024; Zawacki-Richter et al., 2021). These limitations across APOS, RME, and digital learning highlight the need for an instructional model that integrates cognitive theory, contextual learning, and digital affordances into a unified and operational learning design.

ARICT addresses this need by integrating APOS theory, RME principles, and digital learning within a single, structured instructional model articulated through the A–R–I–C–T learning syntax. This syntax operationalizes a constructivist and contextual learning trajectory that guides students from contextual activation and initial responses, through social interpretation and formal construction, to transfer and metacognitive reflection (Agustyaningrum et al., 2022; Zhang et al., 2021). Unlike prior studies that apply APOS, RME, or digital tools separately, ARICT systematically maps each instructional phase to core computational thinking processes, including decomposition, pattern recognition, abstraction, algorithm design, and evaluation (Achmad & Suparman, 2020; Palobo et al., 2025).

The novelty of ARICT lies in positioning mathematics as a structured learning environment for computational thinking development among non-computer-science students through context-rich, digitally supported tasks. In this model, digital learning functions not only as a delivery medium but also as an enabling infrastructure that supports cognitive actions, process development, and schema formation in APOS while simultaneously facilitating contextual exploration and guided reinvention central to RME. By integrating these foundations into a coherent instructional architecture, ARICT advances computational thinking pedagogy beyond fragmented task-based or tool-centered approaches toward a model-driven, theory-grounded framework. Accordingly, this study proposes ARICT as an innovative learning model designed to strengthen students' computational thinking through an integrated instructional architecture. ARICT is conceptualized at the level of a learning model rather than a single method or activity, integrating cognitive construction mechanisms, realistic contextual problem situations, and a digital learning ecosystem that supports exploration, representation, iterative refinement, feedback, and documentation of learning processes. Evidence from prior research suggests that computational thinking–embedded mathematics learning is most effective when supported by structured instructional design and explicit scaffolding (Chytas et al., 2024; Zhang et al., 2025). Figure 1 illustrates the conceptual positioning of ARICT as an integrated learning model that aligns cognitive construction, contextual mathematization, and digital learning environments within computational thinking–oriented mathematics education.

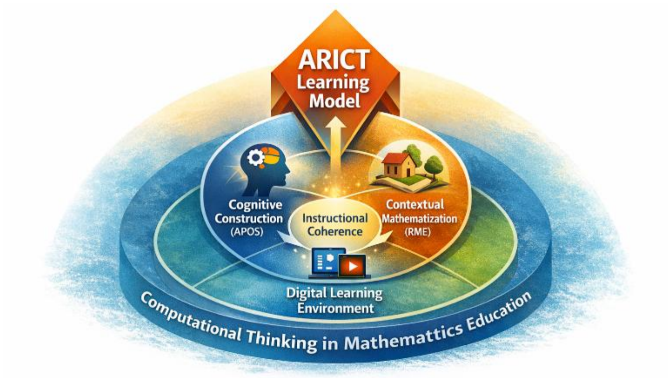

Figure 1 Conceptual Framework of the ARICT Learning Model.

Accordingly, the purpose of this research is to develop and evaluate ARICT in a teacher education context, with particular attention given to whether the model can improve CT components, especially algorithm design and evaluation, while also strengthening mathematical conceptual understanding. This focus responds to contemporary calls for coherent CT training models in teacher education and to the need for stronger operationalization and assessment alignment in CT research in higher education (Darcie et al., 2025; Schoute & Alexander, 2025). This study contributes to mathematics education and educational technology research by offering a model-level instructional innovation that integrates cognitive construction, realistic contextualization, and digital scaffolding into a coherent learning model for CT development and by linking design rationales to assessable outcomes. More specifically, this study pursues the following research objectives:

- To develop the ARICT learning model as a computational thinking oriented instructional framework that integrates APOS theory, realistic mathematics education (RME), and digital learning in the context of vector learning in linear algebra
- To examine the feasibility and validity of the ARICT model through a research and development process; and
- To evaluate the effectiveness of the ARICT model in improving university students' computational thinking and conceptual understanding in vector learning compared with conventional instruction.

This study contributes to mathematics education and educational technology research by offering a model-level instructional innovation that integrates cognitive construction, realistic contextualization, and digital scaffolding into a coherent learning model for CT development and by linking design rationales to assessable outcomes.

2. Materials and methods

2.1. Research design

This study used a research and development (R&D) approach based on the Borg and Gall development model to produce, validate, and evaluate ARICT as an innovative learning model that integrates APOS-informed cognitive constructions, RME-informed contextual mathematization, and digitally mediated learning support to enhance university students' computational thinking (CT) (Hutomo et al., 2022). The Borg and Gall model was selected because it provides a systematic pathway from needs-based product design to iterative field testing and revision, ending with dissemination and implementation as the final R&D objective.

Because the study required both product-quality evidence and effectiveness evidence, the R&D process was implemented within a mixed methods paradigm. Mixed methods are appropriate when qualitative evidence is needed to explain feasibility, implementation challenges, and user responses, whereas quantitative evidence is needed to test the outcome impacts of the developed product (Creswell & Creswell, 2022). In this study, qualitative data supported product revision and interpretation of implementation, whereas quantitative data supported testing the effectiveness of ARICT on CT and conceptual understanding.

2.2. Borg and Gall development stages operationalized for ARICT

The Borg and Gall procedure was operationalized into ten stages adapted to the instructional-model context by Gall in Muliastuti et al., (2023). The research followed a research and development framework adapted from Borg and Gall, as illustrated in Figure 2.

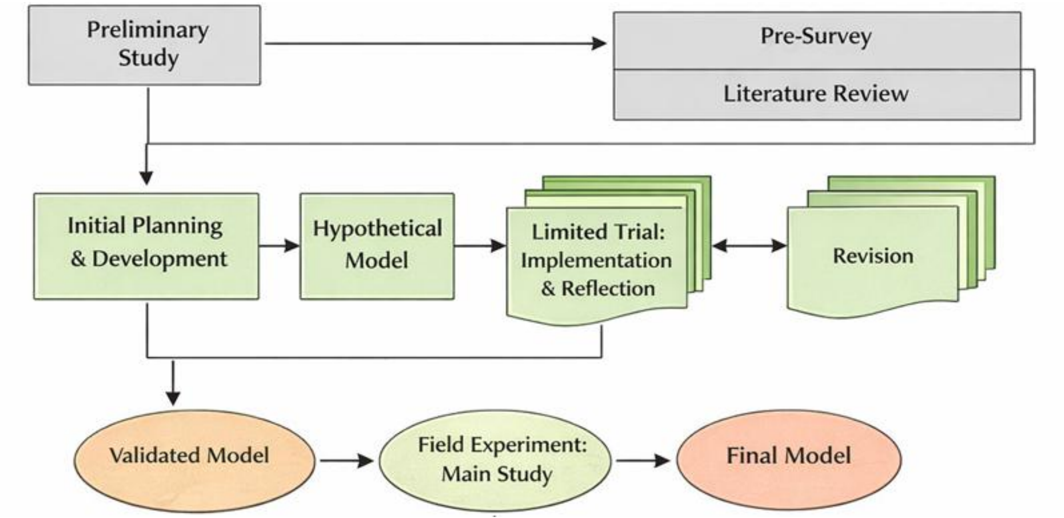

Figure 2 ARICT Research and Development Process.

The process began with research and information collection, which focused on identifying students' computational thinking difficulties, existing instructional practices, and relevant research trends in computational thinking, APOS, RME, and digital learning. Empirical evidence was gathered through surveys, semistructured interviews, classroom observations, and learning-document analysis. In parallel, a systematic evidence scan was conducted via a systematic literature review workflow aligned with PRISMA 2020 principles to enhance transparency and replicability (Page et al., 2021). Bibliometric mapping and thematic clustering were supported via Publish or Perish, Mendeley, and VOSviewer. The findings from this phase informed the planning stage, in which the evidence was translated into a clear product specification for the ARICT learning model. This stage involved defining targeted computational thinking components and indicators, determining learning outcomes and assessment alignment, selecting the content focus, specifying the digital learning infrastructure, and drafting the evaluation design for both limited and main field testing. On the basis of this plan, the study proceeded to develop a preliminary product. A preliminary ARICT package was constructed, consisting of a structured learning syntax aligned with computational thinking indicators, a model-based semester learning plan, digital modules and interactive multimedia, an LMS structure, online and offline student worksheets, and a computational thinking assessment instrument.

The preliminary product was then examined through preliminary field testing. A limited trial was conducted to evaluate the clarity of the learning syntax, the usability of the digital learning flow, the practicality of the materials, and early student learning responses. Data were collected through pretests and posttests of computational thinking, classroom observations, LMS activity traces, student response questionnaires, and lecturer interviews.

Evidence from the limited trial informed the main product revision. Revisions were carried out to improve the internal coherence of the learning syntax, clarify instructional prompts and scaffolds, strengthen feasibility within a digitally supported environment, and refine the computational thinking assessment instrument. Following revision, the model was subjected to main field testing to examine its effectiveness. The revised ARICT package was implemented via a nonequivalent control group pretest–posttest design, which is appropriate for authentic educational settings where random assignment is impractical (Reichardt, 2019).

The intervention was conducted over four weeks, and experimental classes receiving ARICT instruction were compared with control classes receiving conventional instruction while controlling for initial equivalence through pretesting. The control group followed the instructional approach commonly used in the participating program, referred to in this study as conventional instruction. Conventional instruction was primarily lecture-centered and instructor-directed, focusing on explanation of vector concepts, worked examples on the board, and individual problem solving exercises. Digital learning tools or structured computational thinking activities were not systematically integrated in the control classes. Technology use was limited to basic presentation media such as slides or whiteboard explanations and did not involve interactive digital modules, LMS-based learning flows, or computational thinking–oriented tasks. Learning activities in the control classes typically consisted of instructor explanation, guided example solving, and end-of-session exercises drawn from standard course materials.

This instructional format represents the existing pedagogical baseline in the participating institution and therefore served as the counterfactual condition against which the ARICT intervention was evaluated. By clearly distinguishing the ARICT model from the prevailing lecture-based instruction, the study aimed to reduce the risk that the observed effects were attributable merely to technological novelty rather than to the pedagogical design of the ARICT model itself.

On the basis of the results of the main field test, an operational product revision was conducted. This stage focused on refining the implementation guide, strengthening digital activity sequencing, and adjusting the support system to improve the reliability and consistency of enactment across instructional contexts. In line with Borg and Gall's framework, the process includes operational field testing as a subsequent phase aimed at validating the model at a broader scale and under more varied instructional conditions. In the present research program, this phase is positioned as follow-up work to support scaling and sustainability. The development process concludes with final product revision, consolidating the finalized learning model, digital modules, assessment instruments, and observation and practicality tools, followed by dissemination and implementation, which includes journal publications, model documentation, and support for institutional adoption.

2.3. Instrumentation and validation procedures

Instrumentation addresses three evaluation targets: CT outcomes, mathematical conceptual understanding, and implementation fidelity/practicality. The CT instrument targeted five CT components emphasized in this study (decomposition, pattern recognition, abstraction, algorithm design, and evaluation). Conceptual understanding was assessed through a topic-specific test aligned with graph theory learning outcomes. Implementation fidelity was monitored via a structured observation checklist documenting adherence to ARICT phases and key instructional moves.

Expert validation involved mathematics education specialists, APOS–RME experts, and educational technology experts. The validation focused on content validity and construct validity, examining whether the learning syntax and assessment items represented the intended constructs and whether the instructional flow was coherent and feasible in digitally supported learning. For quantitative content validity evidence, expert judgment indices such as Aiken's V can be used; recent methodological work has demonstrated how Aiken's V supports content validity evidence for education instruments on the basis of expert panels (Kania et al., 2024). Revisions were conducted iteratively until the ARICT package and instruments met validity requirements.

2.4. Data sources

This study employed multiple data sources for triangulation. The quantitative data included CT pretest and posttest scores, conceptual understanding pretest and posttest scores, student questionnaire ratings, and descriptive covariates. The qualitative data included classroom observation notes, fidelity checklists, interview transcripts, and learning artifacts produced in online and offline activities. Classroom observations were conducted by independent observers using a structured fidelity checklist designed to capture adherence to the ARICT instructional phases as well as deviations from the intended pedagogical sequence. Digital integration enabled the capture of LMS logs and learning traces as supplementary evidence of participation patterns and learning behaviors. Contemporary learning analytics research emphasizes that log data can enrich the interpretation of learning processes and outcomes while requiring attention to privacy, transparency, and ethical use of learner data (Karaoglan Yilmaz & al., 2022).

2.5. Data Analysis Procedures

The quantitative analysis began with descriptive statistics, assumption checks, and baseline equivalence testing using pretest data. To estimate the effect of ARICT while controlling for baseline CT, posttest outcomes were compared via analysis of covariance (ANCOVA), with pretest scores used as covariates. This approach is commonly used in nonequivalent control group designs to adjust posttest comparisons for preexisting differences and improve interpretability (Ma & Wang, 2023). Effect sizes were reported to quantify practical significance.

3. Results

3.1. Pilot and refinement

Prior to full-scale quasiexperimental implementation, a pilot phase was conducted to examine the feasibility and practicality of the ARICT learning model and to refine the associated instruments and instructional procedures. This phase was designed to determine whether the learning model, assessment tools, and implementation routines could be executed coherently in a realistic classroom context before being scaled to a broader evaluation. In line with feasibility research frameworks, this pilot focused on the clarity of the learning design, student engagement, instructional impact, and operational functionality of the intervention (Bond et al., 2023; Gadke et al., 2021). The feasibility test results, including learning clarity, engagement, and impact, are summarized in Table 1.

Table 1 Results of the feasibility test.

Aspect	Score	Description
Learning Clarity	4.24	The learning material is clearly structured, easy to understand, and systematically presented
Learning Engagement	4.28	The learning activities effectively engage students and encourage active participation
Learning Impact	4.23	The learning material positively influences students' understanding and learning outcomes

As shown in Table 1, all the evaluated aspects achieved high mean scores, indicating that the ARICT learning model was perceived as feasible for classroom use. Learning engagement received the highest score, suggesting that the instructional activities and digital components were effective in promoting active student participation even at the pilot stage. Learning clarity and learning impact were also highly rated, indicating that the students were able to follow the instructional flow and perceived the learning experience as meaningful. Collectively, these results suggest that the ARICT model meets key feasibility criteria related to instructional clarity, learner engagement, and perceived learning value.

Following the feasibility assessment, a broader practicality evaluation was conducted to examine the implementability of the ARICT model under conditions that more closely resembled full classroom deployment. The results of the practicality test are summarized in Table 2.

Table 2 Practicality Test Results.

Aspect	Score	Description
Learning Clarity	4.31	Indicates that learning objectives, instructions, and materials are clearly presented and easily understood by students
Learning Engagement	4.30	Reflects a high level of student involvement, interaction, and motivation during the learning process
Learning Impact	4.32	Demonstrates a strong positive effect of the learning model on students' understanding and learning outcomes

As shown in Table 2, the practicality test results closely mirrored and slightly exceeded those observed in the feasibility phase. All aspects achieved mean scores above 4.30, indicating that ARICT remained clear, engaging, and impactful when implemented at a larger scale. The increase in learning impact suggests that students' perceived benefits became more pronounced as the model was applied more consistently and with refined instructional flow. These findings indicate that ARICT is not only feasible in a pilot context but also practical and scalable for use in regular classroom settings.

Beyond feasibility and practicality ratings, evidence from the pilot phase informed two key categories of refinement. First, the computational thinking (CT) assessment instruments (pretest and posttest) were reviewed to improve clarity and calibration. The item-level response patterns were examined to identify questions that were excessively difficult, overly easy, or ambiguous. On the basis of this analysis, several items were revised or replaced to enhance measurement quality for the main study. This iterative refinement process aligns with contemporary guidelines for instrument development that emphasize expert judgment, empirical testing, and continuous validation as central to building reliable and valid assessment tools (Almanasreh et al., 2022).

Second, classroom observations during the pilot revealed minor instructional flow issues in two phases of the ARICT model. Students benefited from more explicit scaffolding prompts during the Interpret phase, and some participants required additional exploratory time during the Respond phase. These adjustments were incorporated into the revised instructional design to improve accessibility and learning continuity, which is consistent with feasibility frameworks that emphasize user experience, instructional practicality, and implementation constraints as critical dimensions of early-stage intervention research (Liu, 2024).

Minor technical issues were also identified in the GeoGebra-based digital activities, including occasional applet disruptions that interrupted task continuity. These issues were resolved prior to the main implementation. Postpilot interviews further indicated that while students reported high levels of engagement with ARICT, they also experienced increased cognitive challenges. To address this, a clearer orientation and expectation setting were incorporated at the beginning of the intervention, framing the cognitive challenge as an intentional and supportive component of computational thinking development.

3.2. Implementation and data integrity

Building on the pilot findings and refinements reported in Section 3.1, the ARICT learning model was subsequently implemented in a full-scale quasiexperimental study to examine its effectiveness under authentic classroom conditions. Group assignment was conducted at the class level to preserve existing instructional structures. Two classes were assigned to the experimental condition and received the ARICT intervention, whereas the remaining two classes served as the control condition and were taught via conventional instructional approaches. The composition of the research sample, including class allocation, sample size, and instructional approach, is summarized in Table 3.

Table 3 Composition of the Research Sample.

Group	Classes	N	Instruction
Experimental	A, C	41	ARICT
Control	B, D	39	Conventional

The main study was conducted over a four-week intervention period involving four intact classes. To ensure consistent enactment of the instructional design, the ARICT learning model was implemented in the experimental classes by two lecturers who had received prior training on its theoretical foundations and learning phases, thereby minimizing instructor-related variability. The control classes were taught by two experienced lecturers via conventional instructional approaches, primarily interactive board-based instruction followed by guided practice. This parallel instructional arrangement preserved ecological validity while maintaining comparability between groups.

Implementation fidelity was systematically monitored through structured classroom observations supported by a checklist-based protocol. The results indicated high fidelity across both conditions: the ARICT classes consistently enacted the planned instructional phases and digital activity sequences, whereas the control classes maintained conventional teaching patterns. In parallel, data collection was completed with minimal attrition, and all the required quantitative measures of computational thinking and conceptual understanding pretests and posttests were successfully obtained. Qualitative data from observations, student written work, and interview notes were also collected to support triangulation, strengthening confidence that subsequent findings reflect the authentic effects of a well-implemented instructional intervention.

3.3. CT outcomes

Following confirmation of implementation integrity and baseline comparability between groups, the primary analysis examined the effects of the ARICT learning model on students' computational thinking (CT) outcomes. This section reports descriptive patterns of pretest and posttest performance, followed by inferential analyses designed to estimate the intervention effect while accounting for baseline differences. Consistent with best practices in quasiexperimental educational research, both unadjusted descriptive statistics and covariate-adjusted comparisons are presented to provide a transparent and comprehensive account of learning outcomes.

3.3.1. Descriptive Analysis of CT Performance

Descriptive statistics for the CT pretest and posttest scores in both the experimental and control groups are presented in Table 4 to provide an initial overview of learning gains prior to inferential testing.

Table 4 Descriptive Statistics of the CT Pretest and Posttest Scores.

Descriptive Statistics						
	N	Minimum	Maximum	Mean		Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic
Pretest Eksperimen	41	6	25	15.80	.754	4.828
Posttest Eksperimen	41	23	39	32.66	.618	3.960
Pretest Control	39	9	26	15.54	.472	2.946
Posttest Control	39	18	30	24.23	.359	2.241
Valid N (listwise)	39					

As shown in Table 4, both groups demonstrated improvement from pretest to posttest; however, the magnitude of gain was substantially greater in the experimental group. While the mean pretest scores were comparable across groups, the posttest scores indicated a pronounced difference, with students in the ARICT condition achieving markedly higher CT performance than those receiving conventional instruction. These descriptive patterns suggest a strong potential effect of the ARICT intervention, which was further examined via covariate-adjusted analyses.

3.3.2. Assumption Checks for ANCOVA

Prior to conducting inferential analyses, the assumptions underlying the use of ANCOVA were examined, including the normality of score distributions and homogeneity of variances. Normality was assessed via the Shapiro–Wilk test for pretest and posttest CT scores in both the experimental and control groups. The results of these analyses are presented in Table 5.

As shown in Table 5, the Shapiro–Wilk tests indicated that the CT scores in the experimental group did not significantly deviate from normality at either the pretest or the posttest ($p > .05$). In the control group, the posttest scores met the normality assumption, whereas the pretest scores significantly deviated from normality. Given the moderate sample sizes and the robustness of ANCOVA to minor departures from normality, particularly when group sizes are comparable, this deviation was not considered sufficient to invalidate subsequent analyses. The homogeneity of variance was examined via Levene's test, and the results are reported in Table 6.

Table 5 Tests of normality with the Shapiro–Wilk test.

Group	Statistic	df	Sig. (p)
Pretest Eksperimen	0.976	41	0.538
Posttest Eksperimen	0.962	41	0.181
Pretest Control	0.868	39	0.000
Posttest Control	0.952	39	0.099

Table 6 Test of Homogeneity of Variance (Levene's Test).

Test	Levene Statistic	Sig. (p)	Category
Levene (Mean-based)	8.531	0.251	Homogeneous

Levene's test indicated that the assumption of homogeneity of variance was satisfied ($p > .05$), supporting the use of ANCOVA for comparing posttest CT outcomes between groups.

3.3.3. Inferential analysis of CT outcomes

Following the confirmation that the assumptions for ANCOVA were sufficiently met (Tables 5 and 6), an analysis of covariance was conducted to examine the effect of the instructional condition on posttest computational thinking (CT) outcomes while controlling for pretest CT scores. Posttest CT scores were specified as the dependent variable, instructional conditions (ARICT vs. conventional) were specified as the fixed factor, and pretest CT scores were specified as the covariate. To ensure that gains in computational thinking were not achieved at the expense of mathematical content learning, a parallel ANCOVA was also conducted on posttest conceptual understanding scores in vectors in linear algebra, using the corresponding pretest scores as a covariate. The results of these analyses are presented in Table 7.

Table 7 ANCOVA Results on Posttest CT Scores Controlling for the Pretest.

Source	df	F	Sig. (p)	Partial η^2	Effect Size
Pretest (Covariate)	1, 77	2146.657	0.000	0.965	Very large
Class (ARICT vs Control)	1, 77	3663.849	0.000	0.979	Very large

Model fit: $R^2 = 0.987$ (adjusted $R^2 = 0.987$).

As shown in Table 7, the pretest CT score was a highly significant covariate, accounting for a substantial proportion of the variance in the posttest outcomes. More importantly, the instructional condition had a statistically significant and very large effect on posttest CT performance after controlling for baseline differences, $F(1, 77) = 3663.85$, $p < .001$, partial $\eta^2 = .979$, indicating that approximately 97.9% of the variance in adjusted posttest CT scores was attributable to group membership. The exceptionally high model fit ($R^2 = .987$) further suggested that baseline CT and the instructional condition together provided a near-complete explanation of postintervention CT performance. Interpreted within realistic educational research benchmarks, where effect sizes are often attenuated by contextual constraints, these results represent a substantively meaningful instructional impact (Kraft, 2020). In addition, parallel analyses revealed that improvements in CT were accompanied by gains in mathematical content mastery, supporting the interpretation that ARICT strengthened higher-order thinking without diminishing disciplinary understanding.

Overall, the findings indicate that the ARICT learning model produced markedly greater and more consistent improvements in students' computational thinking than did conventional instruction. Although both groups demonstrated learning gains, the magnitude of improvement was substantially greater in the experimental group, and ANCOVA results confirmed that these differences were attributable primarily to the instructional intervention rather than to initial ability variations. At the component level, ARICT enhances all dimensions of computational thinking, with particularly strong gains observed in algorithm design and solution evaluation, areas commonly identified as challenging in complex mathematical problem solving. The structured learning phases and guided digital activities embedded in ARICT appear to provide effective cognitive support for systematic reasoning and reflective evaluation. Taken together, these results position ARICT as a theory-driven, model-level instructional intervention with strong potential for replication and further development in mathematics education and educational technology.

4. Discussion

4.1. Principal contributions of the findings

The purpose of this study was not only to demonstrate that ARICT is feasible but also to evaluate whether a theory-informed, digitally supported learning model can produce meaningful gains in computational thinking (CT) and disciplinary understanding in a higher-education mathematics context. The results provide convergent support for three contributions.

First, ARICT generated meaningful CT improvement relative to conventional instruction under authentic classroom constraints. Second, CT development did not occur at the expense of mathematical learning; rather, the conceptual understanding of vectors also improved. Third, the overall evidence chain, supported by a pilot phase and fidelity monitoring, suggests that the observed impacts are plausibly attributable to the intended learning conditions rather than accidental implementation variation. These contributions respond to persistent concerns in the CT-in-mathematics literature that CT integration often becomes fragmented or tool-centered, limiting consistent learning gains (Namukasa et al., 2023; Nordby et al., 2022).

4.2. Mechanisms Supporting Computational Thinking Improvement

The observed gains in computational thinking (CT) can be interpreted as the result of deliberate alignment between CT indicators and recurring instructional actions embedded within the ARICT learning model. Recent studies consistently emphasize that CT is most likely to develop when instruction makes CT processes explicit, provides repeated opportunities to practice them, and requires learners to externalize, examine, and refine reasoning rather than merely produce final answers (Y. Zhang et al., 2025). This perspective is also consistent with scholarship in mathematics education showing that integrating CT into mathematics is not a straightforward add-on; it requires careful coordination among problem solving, cognitive processes, and instructional transposition to avoid fragmented or tool-centered enactment (Chytas et al., 2024; Van den Heuvel-Panhuizen & Drijvers, 2020). ARICT operationalizes these principles by embedding CT as an explicit learning outcome across its phase-based syntax and leveraging digital learning support to make CT practices more visible, repeatable, and revisable.

Rather than positioning CT as an incidental byproduct of mathematical problem solving, ARICT structures learning activities so that core CT components are systematically cultivated through the A–R–I–C–T sequence. This mechanism is consistent with widely used CT operationalizations that treat CT as multidimensional, commonly including decomposition, pattern recognition, abstraction, algorithmic thinking, and evaluation as central components (Tsai et al., 2021). In ARICT, each phase contributes distinctively to these components, creating a coherent learning trajectory in which CT practices are progressively strengthened across repeated learning cycles.

Decomposition is fostered primarily during the Activate and Respond phases. In Activate, learning begins with a rich and contextualized problem situation that requires students to identify relevant information, constraints, and goals. This early framing encourages students to break down a complex situation into manageable elements by distinguishing what is known, what is unknown, and what must be addressed, which directly corresponds to decomposition as a core CT move (Salac et al., 2021). The response further strengthens decomposition by prompting students to articulate partial ideas and tentative strategies. As learners attempt to respond via their own reasoning, they naturally segment the task into smaller subproblems that can be explored and recombined. Over repeated cycles, this repeated practice helps decomposition become an intentional and habitual thinking practice rather than an implicit skill, aligning with synthesis evidence that CT growth is stronger when learners repeatedly engage in structured CT actions across time rather than in isolated tasks (J. Zhang et al., 2025).

Pattern recognition is most strongly developed during the Respond and Interpret phases. When multiple initial responses are generated, similarities and differences among ideas become visible. Interpret makes this process explicit by structuring peer exchange and collective examination of recurring structures in students' reasoning. This mechanism is theoretically aligned with mathematics-education characterizations of CT, which stresses the importance of cognitive-process work and structured classroom coordination to make disciplinary patterns visible and discussable (Yasin & Nusantara, 2023). This approach is also consistent with research on collaborative computational thinking learning, which emphasizes negotiation, consensus-building, and joint solution construction as pathways through which learners compare strategies and identify recurring structures (Lázaro et al., 2022).

Abstraction is systematically cultivated during the Interpret and Construct phases. Interpreting, informal ideas are reorganized through guided discourse, supporting a shift from contextual details to essential features of the problem and encouraging learners to focus on structure. This refinement is then formalized by introducing representations and generalized forms that encapsulate previously explored reasoning. This gradual transition is consistent with evidence that CT and related higher-order thinking are strengthened when learners are scaffolded to move from informal reasoning toward formal representation rather than being required to adopt formalism prematurely (Dahl-Leonard et al., 2023). It also aligns with CT measurement frameworks that treat abstraction as a distinct, learnable dimension of CT.

Algorithm design is most explicitly developed during the Construct phase. At this stage, students are guided to translate reasoning into systematic, step-by-step procedures that can be executed, tested, and refined. This mechanism is consistent with previous research in mathematics education showing that algorithm construction activities improve students' ability to formalize mathematical reasoning and procedural logic (Demirci & Ergül, 2025; Kovalchuk et al., 2020). Similar findings have also been reported in model-based instructional research, such as problem-based computational modeling and design-based mathematics learning environments, where students are encouraged to iteratively construct and refine solution procedures (Jonsson et al., 2020). This mechanism corresponds directly to algorithmic thinking as a key CT dimension and reflects the view that algorithm design is strengthened when learners actively construct procedures, specify conditions, and refine steps rather than imitate completed solutions. The role of digital support is particularly relevant here because digital environments can increase opportunities for documenting procedures, testing alternatives, and iteratively revising algorithmic representations, a pattern highlighted in recent syntheses of CT pedagogical strategies and assessment practices.

The evaluation is embedded across the Interpret and Transfer and Reflect phases. During Interpret, students critically examine the validity, efficiency, and coherence of alternative strategies through peer discussion and instructor-guided questioning, which strengthens evaluative judgment and the ability to justify decisions. Transfer and Reflect extends evaluation by challenging students to apply constructed procedures in new contexts and to reflect on the effectiveness of their reasoning processes. This emphasis on reflection and monitoring is consistent with evidence that metacognitive scaffolding and systematic reflection can strengthen CT-related dispositions and improve the quality of learners' evaluative and self-regulatory processes (Huang & Chen, 2023). It also aligns with broader findings that CT interventions have stronger and more reliable effects when evaluation and reflection are integrated into the learning cycle rather than being treated as optional after thought.

Across all phases, digital learning functions as an enabling infrastructure rather than the primary driver of instruction. Digital tools support visualization, externalization of reasoning, iterative testing, and documentation of thinking processes. Contemporary reviews indicate that digital environments are most beneficial for CT when they are embedded in coherent pedagogical designs that explicitly guide CT actions and provide opportunities for feedback, iteration, and assessment alignment (Shermin et al., 2019). In ARICT, this support structure plausibly amplifies CT development by making students' reasoning visible and revisable, thereby increasing the frequency and quality of decomposition, pattern recognition, abstraction, algorithm refinement, and evaluation practices. Comparable patterns have also been documented in technology-integrated mathematics learning models such as computational modeling environments and inquiry-based digital mathematics platforms, where digital tools function as cognitive scaffolds that externalize students' reasoning processes and support iterative refinement of problem-solving strategies (Ho et al., 2021). These studies similarly report that CT growth becomes more consistent when digital environments are embedded within coherent instructional designs rather than used as isolated technological tools.

Taken together, these mechanisms explain how ARICT supports CT improvement in a systematic and theoretically grounded manner. The A–R–I–C–T sequence orchestrates a learning trajectory that moves from contextual engagement to structured reasoning and reflective evaluation, with each phase contributing distinctively to specific CT components. This finding is consistent with evidence that CT integration is most effective when implemented as a model-driven instructional system rather than a collection of disconnected activities or tools.

4.3. Mechanisms supporting conceptual understanding gains

A key interpretive value of the findings is the absence of a trade-off between CT development and disciplinary understanding. Similar concerns have been highlighted in prior studies on computational thinking integration in mathematics, where poorly integrated CT activities may lead students to focus on procedural execution rather than conceptual meaning (Lawa et al., 2025; Nordby et al., 2022). However, several intervention studies in mathematics education demonstrate that when CT practices are embedded within conceptually oriented learning cycles, both CT skills and disciplinary understanding can improve simultaneously (Tu & Hsu, 2025). This addresses a common concern that CT integration can dilute conceptual learning when it becomes either tool-driven or overly procedural. In linear algebra, vector topics naturally require representational flexibility, decomposition into components, abstraction to formal operations, and validation of results through geometric or contextual checks. These are cognitively aligned with CT dimensions such as decomposition, abstraction, and evaluation. When instructional cycles require learners to connect representations, justify operations, and verify outcomes, the same routines that strengthen CT can also strengthen conceptual networks.

This interpretation is consistent with scholars who argue that CT in mathematics education must be coordinated with mathematical meaning and not appended as separate activities (Waluyo, 2021). It also aligns with intervention evidence suggesting that technology-supported mathematics learning benefits depend on coherent learning design and explicit scaffolding that ties computational activity to mathematical interpretation (Moore et al., 2024). Therefore, ARICT's impact on conceptual understanding can be read as evidence that the model integrates thinking skill development with content learning rather than treating it as competing goals.

4.4. Methodological interpretation and credibility of claims

Because the study was implemented using intact classes, a nonequivalent control group pretest–posttest design is an appropriate and common approach for higher education contexts where random assignment is not feasible (Lindstromberg, 2025; Ssemugenyi, 2022). The methodological logic of the study rests on three pillars. The first is baseline comparability checks, which reduce (but do not eliminate) selection threats. The second is the use of ANCOVA to adjust posttest comparisons for baseline differences and improve precision in nonrandomized settings (Miyazaki et al., 2022). The third is the inclusion of implementation fidelity evidence, which strengthens internal validity by showing that the intervention was enacted as intended and that the comparison condition did not drift toward ARICT-like practices.

The pilot phase is also methodologically meaningful. Pilot and feasibility work is increasingly viewed as an essential step for improving the usability of procedures, stabilizing digital support, and strengthening instrument quality before a main field test (Bond et al., 2023). In this study, refining item clarity and calibration improved the interpretability of pretest–posttest change as reflecting learning rather than measurement artifacts (Miyazaki et al., 2022; Sun et al., 2020). These methodological decisions collectively reduce avoidable threats to inference quality and are consistent with best practices emphasized in contemporary assessment syntheses.

4.5. Implementation implications and comparative advantages

The findings suggest that ARICT's comparative advantage is not merely the inclusion of digital media but also the way digital learning is subordinated to a coherent pedagogical syntax. This characteristic aligns with design-oriented instructional models frequently used in mathematics education research, such as design-based learning environments and computational modeling frameworks, where digital tools are deliberately integrated within structured pedagogical phases rather than functioning as independent instructional elements (Li et al., 2025) Similar to these models, ARICT organizes learning activities into a sequenced instructional cycle that guides students from contextual problem engagement toward structured reasoning and reflective evaluation . Practically, this implies that implementation should prioritize phase integrity, alignment between tasks and CT indicators, and structured reflection routines as nonnegotiable features. Digital elements such as LMS modules, interactive activities, and online discussion prompts should be treated as part of the support system that makes CT practices observable and revisable rather than as optional add-ons.

High-fidelity enactment in multiple classes indicates that ARICT can be implemented reliably when lecturers are prepared to orchestrate cognitively demanding activities and manage productive struggle. This has direct implications for scalability. Sustainable adoption will require lecturer professional development focusing on facilitation skills, formative feedback, and assessment alignment, as well as institutional support for stable infrastructure and equitable access. Fidelity scholarship emphasizes that fidelity tracking should be feasible and scalable and can function as continuous improvement support rather than only evaluation (Spacciapoli et al., 2022). Systematic reviews also show that clearer fidelity reporting improves the interpretability of intervention effects and supports replication (de Leeuw et al., 2020).

4.6. Limitations and directions for future research

Several limitations should inform interpretation. First, the intervention was implemented over a relatively short period; future work should examine retention and transfer through delayed posttests and tasks requiring CT application in new mathematical topics. Second, the study focused on a single institutional context and a specific topic domain; broader operational testing is needed to establish generalizability across institutions and curricular settings. Third, intact-class designs remain vulnerable to unmeasured confounds even when baseline differences are not statistically significant (Gupta et al., 2023; Ling & Loh, 2023). Future studies can strengthen inference by using matching procedures, adding richer covariates, increasing the number of classes, and modeling lecturer effects more explicitly.

Future research should also explore how CT develops during the ARICT phases. Given the methodological heterogeneity of CT measurement in higher education, triangulating outcome tests with performance-based tasks and processing evidence from solution traces and digital participation patterns would enable stronger mechanism claims (Zhang et al., 2024). Such evidence would also support refinement of phase-specific scaffolds targeting weaker CT components.

4.7. Curriculum and policy implications

At the curriculum level, the findings support the view that CT should be integrated across courses rather than confined to a standalone CT course. Progressive reinforcement across subjects can build a coherent CT pathway, improving the likelihood of transfer beyond single learning. At the institutional level, scalable adoption requires aligning policy and resources with the demands of technology-integrated pedagogy. This includes investing in infrastructure, providing sustained professional development, and recognizing evidence-based instructional innovation and digital pedagogical competence in academic evaluation systems. Overall, the results indicate that ARICT offers a practical mechanism for translating broad goals such as complex problem solving and digital literacy into implementable classroom routines while maintaining, and potentially strengthening, disciplinary understanding.

5. Conclusions

This study demonstrates that ARICT, developed through a Borg and Gall R&D process and implemented with digital learning integration, can substantially enhance students' computational thinking in a linear algebra vector unit while also improving conceptual understanding. The magnitude of the effects on CT and conceptual learning, combined with high implementation fidelity, indicates that the value of ARICT lies in its coherent instructional syntax and its explicit alignment between CT indicators, learning activities, and assessment rather than in technology use alone. For mathematics and teacher education programs, ARICT offers a practical framework for integrating CT as a course-embedded competence and for structuring iterative cycles of reasoning, procedure construction, and verification. Future research should examine retention and transfer through delayed posttests, extend operational field testing to multiple institutions and mathematical topics, triangulate outcome tests with performance-based tasks and process evidence from students' solution traces and digital participation data to strengthen mechanism explanations and scalability guidance.

Acknowledgment

The authors extend their sincere appreciation to all participants in this research project for their time and effort. Their contributions were essential in shaping the research findings.

6. Declarations

6.1. Ethical considerations

The authors have obtained informed consent from all participants in the research for the writing of this article.

6.2. Use of artificial intelligence (AI)

The authors acknowledge the limited use of DeepL solely for minor language refinement, such as grammar checking and improving sentence clarity. The tool was not used in the development of research ideas, data analysis, or interpretation of findings. All intellectual content, conceptualization, and final decisions in this manuscript were fully carried out and validated by the authors, who take complete responsibility for its content. The authors obtained informed consent from all participants in the research for the writing of this article. Full responsibility for the content remains with the authors."

6.3. Conflict of Interest

The authors declare no conflicts of interest.

6.4. Funding

This research did not receive any financial support.

References

Achmad, S. R., & Suparman, S. (2020). Design of e-module with RME approach to improve the creative thinking ability of students. *International Journal of Scientific and Technology Research*, *9*(3), 5228–5233. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85082711127&partnerID=40&md5=48b5256655d897cb5a291fd288eec22a

Agustyaningrum, N., Abadi, A. M., Sari, R. N., & Mahmudi, A. (2022). Students' misconceptions and intuitive reasoning in constructivist mathematics learning environments. *Journal of Mathematical Behavior*, *66*, 100936. https://doi.org/10.1016/j.jmathb.2022.100936

Almanasreh, E., Moles, R. J., & Chen, T. F. (2022). A practical approach to the assessment and quantification of content validity. In *Contemporary Research Methods in Pharmacy and Health Services* (pp. 583–599). Elsevier. https://doi.org/10.1016/B978-0-323-91888-6.00013-2

Bond, C., Lancaster, G. A., Campbell, M., Chan, C., Eddy, S., Hopewell, S., Mellor, K., Thabane, L., & Eldridge, S. (2023). Pilot and feasibility studies: Extending the conceptual framework. *Pilot and Feasibility Studies*, *9*, 24. https://doi.org/10.1186/s40814-023-01233-1

Bozkurt, A., Köse, H., & Akgün, Ö. E. (2022). Integrative approaches to digital learning and educational technologies: A systematic review. *Educational Technology & Society*, *25*(3), 88–102.

Chytas, C., van Borkulo, S. P., Drijvers, P., Barendsen, E., & Tolboom, J. L. J. (2024). Computational thinking in secondary mathematics education with GeoGebra: Insights from an intervention in calculus lessons. *Digital Experiences in Mathematics Education*, *10*, 228–259. https://doi.org/10.1007/s40751-024-00141-0

Creswell, J. W., & Creswell, J. D. (2022). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE.

Dahl-Leonard, K., Hall, C., Capin, P., Solari, E. J., Demchak, A., & Therrien, W. J. (2023). Examining fidelity reporting within studies of foundational reading interventions for elementary students with or at risk for dyslexia. *Annals of Dyslexia*, *73*, 288–313. https://doi.org/10.1007/s11881-023-00279-3

Darcie, I., Gray, R., & Vander Kloet, M. (2025). Sustaining critical minds: How classroom learning cultures shape student thinking dispositions and practices. *Active Learning in Higher Education*, *26*(3 Themed Issue: Active Learning Approaches to Developing Critical Thinking), 543–556. https://doi.org/10.1177/14697874241303508

de Leeuw, R. R., Boer, A. A., & Minnaert, A. E. M. G. (2020). The proof of the intervention is in the implementation: A systematic review about implementation fidelity of classroom-based interventions facilitating social participation of students with social-emotional problems or behavioural difficulties. *International Journal of Educational Research Open*, *1*, 100002. https://doi.org/10.1016/j.ijedro.2020.100002

Demirci, N., & Ergül, E. (2025). Enhancing arithmetic skills and mathematics attitudes through algorithm and coding in primary education. *International Journal of Mathematical Education in Science and Technology*. https://doi.org/10.1080/0020739X.2025.2504555

Dong, W., Li, Y., Sun, L., & Liu, Y. (2024). Developing pre-service teachers' computational thinking: A systematic literature review. *International Journal of Technology and Design Education*, *34*, 191–227. https://doi.org/10.1007/s10798-023-09811-3

Engelbrecht, J., & Borba, M. C. (2024). Recent developments in using digital technology in mathematics education. *ZDM - Mathematics Education*, *56*(2), 281–292. https://doi.org/10.1007/s11858-023-01530-2

Fredriksen, H. (2020). Mathematical modelling and students' meaning making: A realistic mathematics education perspective. *Educational Studies in Mathematics*, *105*(3), 375–393. https://doi.org/10.1007/s10649-020-09957-5

Gadke, D. L., Kratochwill, T. R., & Gettinger, M. (2021). Incorporating feasibility protocols in intervention research. *Journal of School Psychology*, *84*, 1–18.

https://doi.org/10.1016/j.jsp.2020.11.004

Gravemeijer, K. (2020). Rethinking mathematics education: From teaching techniques to learning trajectories. *ZDM–Mathematics Education*, *52*(4), 701–713. https://doi.org/10.1007/s11858-020-01137-9

Gupta, S., Liu, C., Li, S., Chang, F., & Shi, Y. (2023). Association between ability tracking and student's academic and non-academic outcomes: Empirical evidence from junior high schools in rural China. *International Journal of Educational Development*, *103*(June 2021), 102927. https://doi.org/10.1016/j.ijedudev.2023.102927

Herawaty, D., Widada, W., Handayani, S., Febrianti, R., & Anggoro, A. F. D. (2020). Students' obstacles in understanding the properties of the closed sets in terms of the APOS theory. In K. Y.D. & F. of T. T. and E. Universitas Sanata Dharma Mrican, Tromol Pos 29, Yogyakarta (Eds.), *Journal of Physics: Conference Series* (Vol. 1470, Issue 1). Institute of Physics Publishing. https://doi.org/10.1088/1742-6596/1470/1/012068

Ho, W. K., Looi, C. K., Huang, W., Seow, P., Chan, S. W., & Wu, L. (2021). Flowgorithm: A possible passage from algorithmic reasoning to creatively founded mathematical reasoning. In Y. W.-C., M. D.B., & M. M. (Eds.), *Proceedings of the Asian Technology Conference in Mathematics* (pp. 97–111). Mathematics and Technology, LLC. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85127247396&partnerID=40&md5=7fc127b5c51ce7965a00177c5a13a1d0

Huang, W., & Chen, G. (2023). Fostering computational thinking through metacognitive scaffolding and reflective learning strategies. *Computers & Education*, *195*, 104694. https://doi.org/10.1016/j.compedu.2022.104694

Hutomo, B. A., Saptono, S., & Subali, B. (2022). Development of E-module Based on Science, Technology, Engineering, and Mathematics (STEM) To Improve Science Literacy of Junior High School Students. *Journal of Innovative Science Education*, *11*(2), 241–249. https://doi.org/10.15294/jise.v10i1.54066

Jonsson, B., Granberg, C., & Lithner, J. (2020). Gaining Mathematical Understanding: The Effects of Creative Mathematical Reasoning and Cognitive Proficiency. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.574366

Kallia, M., van Borkulo, S. P., Drijvers, P., Barendsen, E., & Tolboom, J. (2021). Characterising computational thinking in mathematics education: A literature-informed Delphi study. *Research in Mathematics Education*, *23*(2), 159–187. https://doi.org/10.1080/14794802.2020.1852104

Kania, N., Kusumah, Y. S., Dahlan, J. A., Nurlaelah, E., Gürbüz, F., & Bonyah, E. (2024). Constructing and providing content validity evidence through the Aiken's V index based on the experts' judgments of the instrument to measure mathematical problem-solving skills. *Research and Evaluation in Education*.

Karaoglan Yilmaz, F. G., & al., et. (2022). Predicting student outcomes using digital logs of learning behaviors: A primer for feature mapping, analytics, and ethical considerations. *Behavior Research Methods*, *54*, 1–22. https://doi.org/10.3758/s13428-022-01939-9

Kovalchuk, M., Voievoda, A., & Prozor, E. (2020). Algorithmic thinking as the meaningful component of cognitive competencies of the future engineer. *Universal Journal of Educational Research*, *8*(11B), 6248–6255. https://doi.org/10.13189/ujer.2020.082263

Lawa, J. A., Madale, V. A., Salic-Hairulla, M. A., Alcopra, A. R., Magsayo, J. R., & Ellare, A. O. (2025). Needs Analysis: Development of a Contextualized Socioscientific Issues (SSI)-Based Module for Enhancing Conceptual Understanding of the Digestive System in Health Science Education. *Seminars in Medical Writing and Education*, *4*. https://doi.org/10.56294/mw2025706

Lázaro, J., Sanz, C., & Moreno, J. (2022). Collaborative problem solving and computational thinking: A systematic analysis of learning processes. *Computers & Education*, *186*, 104545. https://doi.org/10.1016/j.compedu.2022.104545

Li, N., Liu, Z., & Liu, E. (2025). From Conception to Practice: Junior High School Science Teachers' Changes in the Conceptual Change Professional Development. *International Journal of Science and Mathematics Education*. https://doi.org/10.1007/s10763-025-10601-7

Lindstromberg, S. (2025). Eight reasons not to test for baseline group equivalence in a parallel groups pretest-posttest study. *Research Methods in Applied Linguistics*, *4*(3). https://doi.org/10.1016/j.rmal.2025.100254

Ling, M. K. D., & Loh, S. C. (2023). Relationships between cognitive pattern recognition and specific mathematical domains in mathematics education. *International Journal of Mathematical Education in Science and Technology*, *54*(2), 159–179. https://doi.org/10.1080/0020739X.2021.1949059

Liu, T. (2024). Assessing implicit computational thinking in game-based learning: A logical puzzle game study. *British Journal of Educational Technology*, *55*(5), 2357–2382. https://doi.org/https://doi.org/10.1111/bjet.13443

Miyazaki, Y., Kamata, A., Uekawa, K., & Sun, Y. (2022). Bias for Treatment Effect by Measurement Error in Pretest in ANCOVA Analysis. *Educational and Psychological Measurement*, *82*(6), 1130–1152. https://doi.org/10.1177/00131644211068801

Moore, M. L., Pollock, J. R., Elahi, M. A., Smith, J. F., & Rosenow, C. S. (2024). Basic statistical principles. In *Translational Orthopedics* (pp. 117–121). Elsevier. https://doi.org/10.1016/B978-0-323-85663-8.00086-6

Muliastuti, L., Mayuni, I., Nurhaina, A., & Saddhono, K. (2023). Tailoring CEFR to BISOL (Bahasa Indonesia for Speakers of Other Languages) : A Model for Integrative Language Teaching Materials. *International Journal of Language Education*, *7*(4), 590–601. https://doi.org/10.26858/ijole.v7i4.53219

Namukasa, I. K., Gecu-Parmaksiz, Z., Hughes, J., & Scucuglia, R. (2023). Technology maker practices in mathematics learning in STEM contexts: a case in Brazil and two cases in Canada. *ZDM – Mathematics Education*, *55*(7), 1331–1350. https://doi.org/10.1007/s11858-023-01534-y

Nordby, S. K., Bjerke, A. H., & Mifsud, L. (2022). Computational Thinking in the Primary Mathematics Classroom: a Systematic Review. *Digital Experiences in Mathematics Education*, *8*(1), 27–49. https://doi.org/10.1007/s40751-022-00102-5

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., & al., et. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71

Palobo, M., Sulaiman, R., & Rahaju, E. B. (2025). Construction of the Area of Rectangle Concept from the Perspective of APOS Theory. *TEM Journal*, *14*(1), 149–159. https://doi.org/10.18421/TEM141-14

Rao, T. S. S., & Bhagat, K. K. (2024). Computational thinking for the digital age: A systematic review of tools, pedagogical strategies, and assessment practices. *Educational Technology Research and Development*, *72*, 1893–1924. https://doi.org/10.1007/s11423-024-10364-y

Reichardt, C. S. (2019). *Quasi-experimental design BT - SAGE research methods foundations* (P. Atkinson, S. Delamont, A. Cernat, J. W. Sakshaug, & R. A. Williams (eds.)). SAGE. https://doi.org/10.4135/9781526421036755088

Salac, J., Thomas, C., Butler, C., & Franklin, D. (2021). Investigating the Role of Cognitive Abilities in Computational Thinking for Young Learners. *ICER 2021 - Proceedings of the 17th ACM Conference on International Computing Education Research*, 2–17. https://doi.org/10.1145/3446871.3469746

Schoute, E. C., & Alexander, P. A. (2025). A Critical Analysis of Critical Thinking Interventions in Higher Education. *Review of Educational Research*. https://doi.org/10.3102/00346543251352539

Serpe, A., & Frassia, M. G. (2020). Task Mathematical Modelling Design in a Dynamic Geometry Environment: Archimedean Spiral's Algorithm. In Y. D. Sergeyev,

D. E. Kvasov, Y. D. Sergeyev, & D. E. Kvasov (Eds.), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 11973 LNCS* (pp. 478–491). Springer. https://doi.org/10.1007/978-3-030-39081-5_41

Shermin, A., De Silva, B., & Wesumperuma, A. (2019). Groupwork Assessment Development (GAD): A framework for developing an effective group work assessment. *ACIS 2019 Proceedings - 30th Australasian Conference on Information Systems*, 851–861. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85173630551&partnerID=40&md5=d9b03969046fee20b0841e32a3a43d48

Spacciapoli, M., Viana, M., Saunders Wilder, O., Sullivan, J., McCallum, T., & Wilder-Smith, B. (2022). An equitable and scalable approach to track fidelity of implementation in partnership with teachers. *Frontiers in Education*, *7*, 1020204. https://doi.org/10.3389/feduc.2022.1020204

Ssemugenyi, F. (2022). Trapped at the crossroads: Does problem-based learning make a difference? The moderating role of traditional mode of instruction. *Cogent Education*, *9*(1). https://doi.org/10.1080/2331186X.2022.2068398

Sun, X., Liu, Y., Xin, T., & Song, N. (2020). The Impact of Item Calibration Error on Variable-Length Cognitive Diagnostic Computerized Adaptive Testing. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.575141

Tsai, M.-J., Liang, J.-C., Lee, S. W.-Y., & Hsu, C.-Y. (2021). Structural validation for the developmental model of computational thinking. *Journal of Educational Computing Research*, *59*(2), 300–325. https://doi.org/10.1177/07356331211017794

Tu, S.-F., & Hsu, H.-W. (2025). The Analysis of Teachers' Corrective Experience on Receiving School Counselors' Consultations. *Bulletin of Educational Psychology*, *56*(4), 917–946. https://doi.org/10.6251/BEP.202506_56(4).0006

Van den Heuvel-Panhuizen, M., & Drijvers, P. (2020). *Realistic mathematics education BT - Encyclopedia of mathematics education* (S. Lerman (ed.); pp. 521–525). Springer. https://doi.org/10.1007/978-3-030-15789-0_170

Waluyo, A. A. (2021). Critical thinking skills assessment instrument in physics subjects: How to develop a four tier diagnostic test? *Journal of Physics: Conference Series*, *1918*(2). https://doi.org/10.1088/1742-6596/1918/2/022011

Yasin, M., & Nusantara, T. (2023). Characteristics of Pattern Recognition to Solve Mathematics Problems in Computational Thinking. In H. Habiddin & N. Farida (Eds.), *AIP Conference Proceedings* (Vol. 2569). American Institute of Physics Inc. https://doi.org/10.1063/5.0112171

Zawacki-Richter, O., Alturki, U., Aldraiweesh, A., & Aljahdali, A. (2021). Mapping research in digital learning environments: A systematic review of literature. *The Internet and Higher Education*, *50*, 100805. https://doi.org/10.1016/j.iheduc.2021.100805

Zhang, J., Wu, Y., Ning, Y., & Shi, Y. (2025). From evidence to insight: An umbrella review of computational thinking research syntheses. *Journal of Intelligence*, *13*(12), 157. https://doi.org/10.3390/jintelligence13120157

Zhang, L., Chen, G., & Sun, D. (2021). The role of metacognitive reflection in developing computational thinking. *Computers & Education*, *168*, 104193. https://doi.org/10.1016/j.compedu.2021.104193

Zhang, X., Aivaloglou, F., & Specht, M. (2024). A systematic umbrella review on computational thinking assessment in higher education. *European Journal of STEM Education*, *9*(1), 2. https://doi.org/10.20897/ejsteme/14175

Zhang, Y., Wu, X., Ning, Y., & Shi, Y. (2025). Effectiveness of computational thinking interventions: An umbrella review of meta-analyses. *Review of Educational Research*, *95*(1), 1–38. https://doi.org/10.3102/00346543241234567