REVIJA ZA ELEMENTARNO IZOBRAŽEVANJE
JOURNAL OF ELEMENTARY EDUCATION
Vol. 19, No. 2, pp. 229–246, June 2026

Critical Thinking in Elementary Mathematics: A Mindful, Meaningful, and Joyful Learning Model

Pariang Sonang Siregar[1], Neni Hermita[2], Firman Santosa[3] & Tety Herti[1]
[1]Univeristas Rokania, Faculty of Educational Sciences, Rokan Hulu, Republic of Indonesia
[2]Universitas Riau, Faculty of Teacher Training and Education, Pekanbaru, Republic of Indonesia
[3]Universitas Rokania, Faculty of Computer Science, Rokan Hulu, Republic of Indonesia

Korespondenčni avtor/Corresponding author
pariangsonangsiregar@gmail.com

Potrjeno/Accepted
6. 4. 2026

Objavljeno/Published
30. 6. 2026

Abstract/Izvleček
This study evaluates a *deeper learning* model integrating mindful, meaningful, and joyful learning to bridge the procedural-conceptual knowledge gap. Using a sequential exploratory mixed-method design within a Lesson Study framework, we involved 7 teachers and 20 Indonesian fourth graders. Quantitative results showed significant critical thinking improvement (F = 528.603, p < .000) with a large effect size (d = 7.36). Qualitative findings revealed that the model established a psychologically safe environment, facilitating student-led knowledge construction. Despite the small sample, this study provides a validated pedagogical framework for fostering higher-order mathematical reasoning through affective safety and metacognition.

Keywords:
Deeper learning, pedagogical model, critical thinking, mathematics education, elementary school.

Kritično mišljenje pri osnovnošolski matematiki: čuječ, osmislitven in radosten model učenja
Študija ocenjuje model *poglobljenega učenja*, ki vključuje čuječe, osmislitveno in radostno učenje za premostitev vrzeli med proceduralnim in konceptualnim znanjem. Z uporabo mešanih metod v okviru študija lekcij je sodelovalo 7 učiteljev in 20 indonezijskih četrtošolcev. Kvantitativni rezultati so pokazali značilno izboljšanje kritičnega mišljenja (F = 528,603, p < 0,000) z izjemno velikim učinkom (d = 7,36). Kvalitativne ugotovitve razkrivajo, da je model vzpostavil psihološko varno okolje, ki spodbuja samostojno konstrukcijo znanja. Kljub majhnemu vzorcu študija ponuja potrjen pedagoški okvir za spodbujanje matematičnega sklepanja višjega reda prek afektivne varnosti in metakognicije.

Ključne besede:
poglobljeno učenje, proceduralno znanje, konceptualno znanje, kritično mišljenje, poučevanje matematike, osnovna šola.

UDK/UDC
37.091.3:51:159.955

DOI https://doi.org/10.18690/rei.5575

Besedilo / Text © 2026Avtor(ji) / The Author(s)
To delo je objavljeno pod licenco Creative Commons CC BY Priznanje avtorstva 4.0 Mednarodna. Uporabnikom je dovoljeno tako nekomercialno kot tudi komercialno reproduciranje, distribuiranje, dajanje v najem, javna priobčitev in predelava avtorskega dela, pod pogojem, da navedejo avtorja izvirnega dela. (https://creativecommons.org/licenses/by/4.0/).

University of Maribor Press

Introduction

A significant body of international research highlights a persistent gap between the educational aim of fostering critical thinking and the prevalent classroom practices in mathematics (Stein, Grover, and Henningsen 2006; Wang and Abdullah, 2024). In many educational contexts, mathematics instruction remains heavily focused on procedural fluency, rote memorisation, and the pursuit of a single correct answer (Foster, 2018; Li and Wen, 2023). This focus often comes at the expense of conceptual understanding, creating a barrier to the development of higher-order cognitive skills (Rittle-Johnson and Alibali, 1999).

Critically, this gap is not merely a failure of memory, but a complex disconnects between different types of mathematical knowledge. As Haapasalo and Kadijevich (2000) argue, procedural knowledge (knowing "how") and conceptual knowledge (knowing "why") are not opposing forces but are deeply interdependent and mutually reinforcing. Factual and procedural knowledge provide the algorithmic efficiency necessary for problem-solving, while conceptual knowledge provides the logical foundation and verbalisation required for knowledge transfer to novel situations (Levin, 2018). A balanced mathematics education, therefore, does not reject memorisation of facts but seeks a cognitive equilibrium where procedures are illuminated by conceptual meaning. However, a "knowing-doing gap" often persists while teachers value critical thinking, they frequently lack practical, operationalised models to facilitate this balance in a way that moves students beyond rote drills toward genuine conceptual engagement (Kannadass et al., 2023).

This study uses the term *deeper learning* in its specific pedagogical sense (Fullan, 2016; McEachen, Fullan, and Quinn 2018), referring to a process where learners actively construct knowledge, make connections, and develop 21st-century competences. We operationalise this concept through a synergistic model built on three core principles. First, mindful learning is defined as a metacognitive state where students are present, aware, and actively regulating their reasoning strategies rather than following steps blindly (Kudesia, 2019; Jankowski and Holas, 2014). This is integrated with meaningful learning, a constructivist process where abstract symbols and procedural facts are anchored in real-world contexts to ensure that new information is coherently integrated into existing cognitive structures (Jonassen and Strobel, 2006; Usiskin, 1982; Nguyen, 2025). Finally, the model incorporates joyful learning, an affective condition characterized by psychological safety, high engagement, and the reduction of math-related anxiety, which often acts as a physiological barrier to the complex cognition required for deep understanding (Clanton Harpine, 2024; Mystakidis et al., 2019; Sablić, Mirosavljević, and Marinac 2025).

The Indonesian education system serves as a pertinent case study of this global challenge. In Riau Province, our initial field analysis indicated that 67.5% of students possessed low critical thinking abilities, largely because instruction prioritized curriculum completion over the deep integration of procedural and conceptual knowledge (Sarwanto, Fajari, and Chumdari, 2021). The novelty of this research lies in the deliberate synthesis of these three principles into a cohesive, practical model. Unlike previous student-centred approaches that may focus heavily on context (*meaningful*) or activity (*mindful*), our model explicitly positions the affective (*joyful*) as the prerequisite "gateway" that lowers the amygdala's anxiety response, thereby unlocking the higher-order (Mindful) metacognition necessary to process (Meaningful) contextual tasks.

In doing so, the research makes a distinct contribution to the field. While much literature discusses the theory of *deeper learning*, this paper addresses the practical "knowing-doing gap" by providing a validated, operationalised pedagogical framework that demonstrates how affective safety and metacognition can be balanced with mathematical rigour. Furthermore, by providing a rich, mixed-methods case study from an Indonesian context, this research offers empirical evidence and transferable insights for educators internationally, aligning with the need for global perspectives on bridging the procedural-conceptual divide in mathematics.

Consequently, this research seeks to answer the following two questions:

RQ1: How do teachers and students experience the implementation of a mathematics model based on mindful, meaningful, and joyful principles?

RQ2: What is the effect of this pedagogical model on the critical thinking skills of elementary school students?

Literature Review

The iterative nature of conceptual and procedural knowledge

A long-standing debate in mathematics education concerns the relationship between procedural and conceptual knowledge. Current educational literature emphasizes that these two are not mutually exclusive; instead, they exist in a bidirectional and iterative relationship (Rittle-Johnson and Alibali, 1999). Procedural knowledge involves the ability to execute action sequences (algorithms) to solve problems, while conceptual knowledge involves an integrated and functional grasp of mathematical ideas (Haapasalo and Kadijevich, 2000).

Critical thinking is inhibited when these two domains are bifurcated. When students learn procedures without conceptual anchoring, they often produce "correct" answers without the ability to justify their reasoning or adapt to novel problems (Richards, Hayes, and Schwartzstein 2020). Conversely, conceptual understanding without procedural fluency can lead to high cognitive load during complex tasks. Our model posits that critical thinking emerges when students are encouraged to verbalise their conceptual understanding during procedural execution. This verbalisation acts as a bridge, transforming a "rote" procedure into a "meaningful" cognitive action (Chen, Chen and Lai, 2025).

The affective dimension: Joyful learning as a cognitive prerequisite

Traditional mathematical instruction often neglects the affective domain, focusing exclusively on cognitive outcomes. However, the role of *Joyful learning* is grounded in the *affective filter* hypothesis, suggesting that negative emotions like anxiety or boredom create a physiological blockage in the brain's prefrontal cortex (Willis, 2007). In mathematics, anxiety is a pervasive barrier that restricts working memory capacity, making critical thinking nearly impossible (Sablić, Mirosavljević, and Marinac, 2025). By operationalising "joy" as psychological safety and curiosity-driven exploration, our model creates a *low-filter* environment. This safety is not a trivial addition of fun but a deliberate pedagogical strategy to unlock higher-order thinking. When students are in a joyful state, they are more likely to engage in intellectual risk-taking: the willingness to propose hypotheses and test logical ideas without the paralyzing fear of being incorrect (Clanton Harpine, 2024).

Metacognition and Mindful Engagement

Mindful learning in our model is synonymous with active metacognition—the ability of students to monitor, direct, and evaluate their own thinking processes. According to McEachen, Fullan, and Quinn (2018), *deeper learning* is only achieved when students move beyond being receivers of information to becoming directors of their learning. This requires *mindfulness*, a state of cognitive presence where students consciously choose strategies rather than defaulting to mechanical repetition.

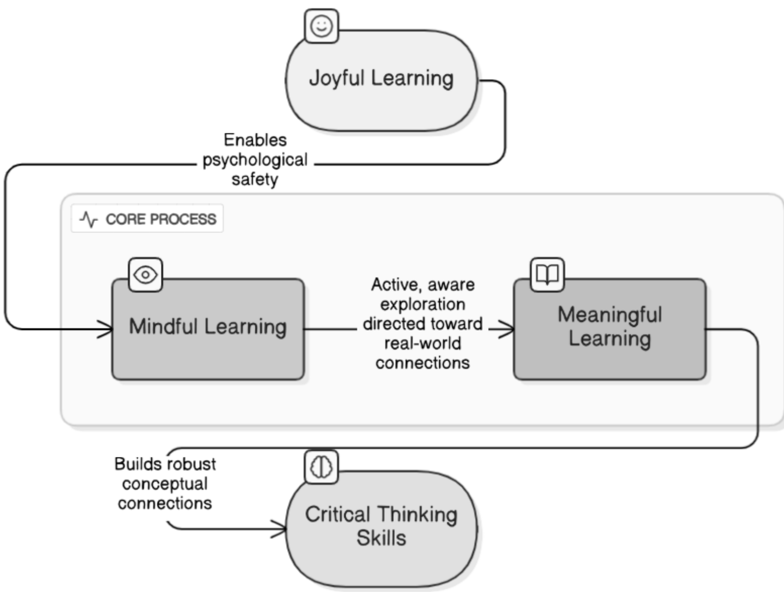
Critical thinking indicators, such as those identified by Ennis (1989), including inference, deduction, and evaluation, are inherently metacognitive. For a student to evaluate a geometric problem, they must first be *mindfully* aware of the characteristics of the shapes involved and the limitations of their chosen formula (Jonassen and Strobel, 2006). Our model facilitates this through "What if?" questioning strategies, pushing students to reflect on the process of discovery rather than just the final answer.

Positioning against existing integrated models

While *Realistic Mathematics Education* (RME) emphasizes the *meaningful* through contextualization and the use of real-world models (Van Den Heuvel-Panhuizen, 2003), it sometimes under-theorizes the role of the teacher in managing the affective environment. Similarly, gamification focuses on the *joyful* but can sometimes sacrifice mathematical depth for edutainment (Beattie, 2024). Our model differs by theorizing a specific, non-linear synergy, as illustrated in Figure 1: *Joyful learning* functions as an affective gateway that lowers the cognitive load created by anxiety, allowing for *mindful* metacognition to be applied to solve *meaningful* real-world problems. This synthesis ensures that the affective, metacognitive, and constructivist elements work as an interdependent ecosystem rather than additive parts.

Figure 1.

Synergistic Framework. The overlaps signify that Joyful learning provides the emotional stability for Mindful focus, while Mindful metacognition provides the focus required to construct Meaning from context.



Method

We employed a sequential exploratory mixed-method design with a one-group pre-test-post-test configuration (Creswell and Creswell, 2009). The study lasted two months, consisting of ten intensive instructional sessions designed through a lesson study (open class) format. This format followed a systematic three-stage cycle (*Plan-Do-See*) to refine the instructional intervention (Thinwiangthong et al., 2021).

Instructional intervention: The lesson study cycle

To operationalise the mindful, meaningful, and joyful principles, the intervention was structured as follows:

1. *Plan.* The group of 7 teachers met for three 2-hour collaborative sessions prior to each open class. They analysed specific student conceptual difficulties (e.g., confusing perimeter with area) and designed *Meaningful* contextual tasks using concrete media like origami and classroom furniture. The lessons were designed to prompt *Mindful* reflection through specific questioning prompts.
2. *Do (Open Class).* The model teacher implemented the designed lesson plan in the fourth-grade class. During these sessions, the other six teachers acted as observers, strategically positioned to record student dialogue, body language (affective markers for *Joyful*), and problem-solving steps rather than focusing on the teacher's delivery.
3. *See (Reflection).* Immediately following the lesson, the team engaged in structured reflection. Observers provided evidence-based feedback on student engagement and misconceptions. For example, if students struggled to transition from origami manipulation to formula abstraction, the next "Plan" stage was adjusted to include more bridging activities.

Participants and selection

The study involved one intact class of 20 fourth-grade students (aged 9-10) and a focus group of seven educators. Teachers were purposively selected from a local cluster based on their history of pedagogical innovation (Table 1). Categorization was based on preliminary peer-interviews regarding their use of project-based and student-centred methods.

Table 1
Demographic profiles of teacher participants (n=7)

Teacher	Gender	Innovation category	Teaching Experience	Role in study
Teacher 1	Male	Always	15 years	Observer
Teacher 2	Female	Often	12 years	Observer
Teacher 3	Female	Often	8 years	Observer
Teacher 4	Female	Rarely	20 years	Model teacher
Teacher 5	Female	Always	5 years	Observer
Teacher 6	Male	Often	10 years	Observer
Teacher 7	Female	Always	9 years	Observer

To ensure that the intervention focused on critical thinking transitions rather than basic arithmetic remediation, it was verified prior to the study that all students had achieved a baseline competency in single-digit multiplication and 2D shape identification. Furthermore, to maintain instructional validity and reflect standard classroom conditions, only one primary teacher conducted the instruction, despite the collaborative involvement of seven teachers in the lesson design.

Data collection and instruments

Two researchers with doctoral-level training in mathematics education conducted the observations. They underwent a 10-hour synchronization training to ensure high inter-rater reliability on the 4-point observation rubric (Appendix C). Data were gathered using a combination of a mathematical test and semi-structured interviews. The 7-item open-ended mathematical test focused on Geometry, specifically the surface area of quadrilaterals. This domain was selected as a proof-of-concept because its visual and spatial characteristics facilitate the transition from concrete manipulation to abstract deduction. Validation of the instrument involved both content and construct validity. Content validity was established through expert review by three specialists in mathematics education. Construct validity was subsequently assessed via Confirmatory Factor Analysis (CFA), yielding favourable results with an Average Variance Extracted (AVE) of 0.57 and a Heterotrait-Monotrait (HTMT) ratio of 0.500. Reliability was further confirmed with a Composite Reliability (CR) of 0.762 and a Cronbach's Alpha of 0.810 (Cronbach, 1951; Kline, 2013). The specific mapping of items to critical thinking indicators and cognitive levels is detailed in Table 2, while a sample item from this instrument is provided in Appendix A.

Table 2
Mapping of test items to critical thinking indicators and taxonomy levels

Indicators (Ennis, 1989)	Item Example	Taxonomy Level
Elementary Clarification	Calculate remaining land area after a portion is used for a garden.	C4 (Analysis)
	How much land area can Dino use for a playground? (Evaluating constraints)	C5 (Evaluation)
Building Basic Concept	How do you calculate the area of a parallelogram-shaped staircase? (Applying concept)	C6 (Creation)
Inference	Jiko's kite is 150 cm^2. Dino's is twice as large. Calculate Dino's kite circumference.	C5 (Evaluation)
Advanced Clarification	Two shapes have areas of 10 m^2 and 4 m^2. Name two types of flat shapes.	C4 (Analysis)
	Determine the possible dimensions of each side of the two flat shapes.	C6 (Creation)
Strategies and Tactics	A blackboard has a circumference of 280 cm^2. Calculate the size of the side.	C6 (Creation)

Scoring reliability was established through an inter-rater process (Cole, 2024), resulting in a high Cohen's Kappa of 0.87. Additionally, semi-structured interviews were conducted with all seven participating teachers and six students, selected through stratified purposive sampling based on achievement levels. These sessions were guided by a validated protocol (Kallio et al., 2016) to ensure systematic yet flexible exploration of the participants' implementation experiences.

Trustworthiness and triangulation

Trustworthiness was ensured through data triangulation and member checking (Ang et al., 2016). To strengthen the credibility of the qualitative coding, the two researchers co-coded a 20% sample of the transcripts, resolving discrepancies through consensus meetings until an agreement rate of 90% was achieved.

Data analysis

Data analysis followed a comprehensive approach to synthesize qualitative and quantitative findings (Creswell and Creswell, 2009; Fitrah et al., 2024). Qualitative data were subjected to a hybrid thematic analysis (Braun and Clarke, 2006) managed via Atlas.ti 24 software (Vila-Henninger, 2019). This hybrid approach combined deductive coding, based on the *a priori* framework of mindful, meaningful, and joyful-learning, with inductive coding to capture emerging strategies and challenges directly from participant voices. The process involved six recursive phases: data familiarization, initial code generation, searching for themes, reviewing themes, defining theme names, and producing the final report (Braun and Clarke, 2006).

Quantitative scores were analysed using R version 4.5.1 (Chambers, 2008). Prior to inferential testing, data were screened for normality using the Shapiro-Wilk test to ensure the appropriateness of the parametric procedures (Shapiro and Wilk, 1965). Descriptive statistics (Mean and Standard Deviation) were calculated to assess the general growth in critical thinking scores. To determine the statistical significance of the improvement, a paired-samples t-test was conducted, with results reported as an F-statistic ($F = t^2$) to align with broader variance analysis conventions (Polatbekova et al., 2025). Furthermore, Cohen's *d* was calculated to evaluate the effect size and practical significance of the pedagogical intervention, independent of the small sample size (Cole, 2024).

Findings

RQ1: Implementation fidelity and participant experiences

The implementation of the model was characterized by high levels of teacher fidelity and deep student engagement across all three core principles. Observational data (Table 3) confirmed that the lesson designs were translated effectively into the classroom environment.

Table 3

Teacher observation scores on model implementation (n=7 teachers)

Indicators	T1	T2	T3	T4	T5	T6	T7	Average	Category
Mindful learning	82.5	92.5	80	87.5	67.5	77.5	97.5	87.57	Very good
Meaningful learning	77.5	92.5	90	87.5	87.5	82.5	92.5	87.14	Very good
Joyful learning	70	82.5	90	92.5	95	85	77.5	84.64	Very good

Note. Scores represent the percentage of positive indicators observed for each teacher across all sessions

The thematic analysis revealed that the model facilitated a profound shift in classroom culture. Table 4 provides a summary of the emergent themes, supported by representative participant voices that illustrate the transition from surface to *deeper learning*.

Table 4
Summary of qualitative themes from teacher and student interviews

Theme	Sub-themes: Teacher perspective (strategies)	Sub-themes: Student perspective (experiences)
Mindful learning	Fostering student-led discovery: "Providing open-ended questions and assignments to encourage exploration." (Teacher 5)	Developing metacognitive strategies: "Actively observing the environment to confirm or disconfirm hypotheses." (Student 6, high)
	Using concrete media: "Employing puzzles, origami, and videos as tools for students to build formal concepts." (Teacher 5)	Strategic use of resources: "Using videos and classroom objects to find answers independently." (Student 2, low)
	Focus on reasoning: "Asking students *how* they got an answer, not just *if* it was correct." (Teacher 7)	Building conceptual links: "Using characteristics of shapes to test formulas on new objects." (Student 6, high)
Meaningful learning	Connecting to real-world contexts: "Inviting students to use their own life experiences as the basis for problems." (Teacher 3)	Recognising math in daily life: "Seeing geometry 'in class, at home, and everywhere'." (Student 1, low)
	Using contextual problems: "Symbolising contextual problems into formal mathematical language." (Teacher 4)	Identifying practical application: "Understanding that math (e.g., multiplication) is used for real tasks like buying things." (Student 3, int)
	Simulating real life: "Using role-playing (e.g., shopping) to connect math operations to tangible actions." (Teacher 1)	Transferring knowledge: "Actively applying classroom learning at home (e.g., 'measure the surface of a table')." (Student 5, high)
Joyful Learning	Creating psychological safety: "Giving freedom to express logical ideas *without* fear of being 'blamed' for errors." (Teacher 1)	Increased affective engagement: "Feeling 'enthusiastic' and 'never giving up'." (Student 6, high)
	Gamification and play: "Using games, puzzles, and role-playing to frame learning activities." (Teacher 2)	Learning as play: "Describing learning in terms of 'games', 'puzzles', and 'exploring'." (Student 4, int)
	Valuing the process: "Treating incorrect answers as part of the 'process of discovery', not as failures." (Teacher 1)	Intrinsic motivation: "Being motivated by the activity itself (e.g., finding answers, playing games)." (Student 2, low)

Addressing affective barriers in low-achieving students

A critical finding related to the *Joyful* component was its impact on students who initially struggled with multiplication. Rather than feeling discouraged by procedural speed, these students found the contextual manipulation (e.g., tiling physical quad rilaterals) provided a "concrete bridge" that reduced their anxiety. Teacher 7 observed: *"I saw students who usually hide during math class suddenly leading their groups because they 'saw' the geometry through the origami tasks."*

This suggests that the model's affective safety is particularly vital for those with lower baseline confidence, addressing a key concern about the feasibility of higher-order thinking for all learners.

The human dimension of pedagogical transition

While fidelity was high, interviews highlighted the internal challenges faced by teachers. Teacher 4, who was categorized as 'Rarely' innovative, admitted: *"It was difficult to trust the students to lead. Initially, I felt 'naked' without my whiteboard notes. I had to resist the urge to just tell them the formula to save time."*

This suggests that the *Meaningful* transition requires not just new materials, but a fundamental shift in professional identity from a "provider of answers" to a "facilitator of discovery."

RQ2: Quantitative effectiveness of critical thinking

The quantitative results confirmed that the qualitative shifts in engagement were matched by significant cognitive gains, as summarized in Table 5 and Table 6. Preliminary assumption testing using the Shapiro-Wilk test confirmed that the difference scores were normally distributed ($W = 0.98$, $p > 0.05$), justifying the use of parametric testing. The analysis of pre-test and post-test scores showed a robust improvement across all critical thinking indicators.

Table 5

ANOVA statistics for critical thinking scores (pre-test and post-test)

Source	*Sum of squares*	*df*	*Mean square*	*F*	*Sig.*
Model Intervention	1416.100	1	1416.100	528.603	0.000
Residual (Within groups)	101.800	38	2.679		

Note. The "Between Groups" reflects the difference between the pre-test and post-test means. The F-statistic is significant at the .05 level.

Table 6

Descriptive statistics for critical thinking scores (n=20)

Test phase	Mean (M)	Std. Deviation (SD)
Pre-test	9.1	1.94
Post-test	21	1.26

The difference between pre-test and post-test means was statistically significant ($F = 528.603$, $p < .000$). To determine if this significance was merely a result of the sample size or had practical value, Cohen's d was calculated at 7.36. This value indicates an extraordinarily large effect size, suggesting that the *Deeper Learning* intervention accounted for a substantial portion of the variance in student achievement.
Analysis of specific test items revealed that the greatest gains occurred in Item 7 (Evaluation/Creation). In the pre-test, only 10% of students could model a real-world geometric problem; in the post-test, this increased to 85%. Student A's response in Appendix B exemplifies this growth, showing not just a correct numerical answer, but a logical justification based on the optimization of spatial properties.

Discussion

Integrating procedural and conceptual knowledge through verbalisation
The primary goal of this study was to evaluate a model that bridges the disconnect between procedural fluency and conceptual understanding. The findings confirm that when students are mindfully engaged in meaningful tasks, they do not merely execute algorithms; they construct a cognitive map of the underlying mathematical relations. This aligns with the iterative theory of Rittle-Johnson & Alibali (1999), suggesting that conceptual knowledge drives procedural gains. Crucially, the qualitative data showed that the *Meaningfu*l context (e.g., floor tiling) provided the necessary anchor for students to verbalise their reasoning. As suggested by Haapasalo & Kadijevich (2000), this verbalisation is the key mechanism that transforms a rote procedure into a conceptual tool, allowing for the high-order transfer observed in the post-test creations.

The synergy of affect and cognition: Joy as a prerequisite
A significant contribution of this research is the operationalisation of the *Joyful* principle not as a leisure activity, but as a critical cognitive prerequisite. Reviewer 2 correctly questioned the feasibility of joy for students struggling with prerequisites. However, our findings suggest that for these students, the "blame-free" and exploratory nature of the model was more impactful than for high-achievers. By lowering the "affective filter" (Willis, 2007), we observed that low-achieving students who previously hid during math class became active participants. This confirms that affective safety acts as a physiological gateway to the prefrontal cortex, enabling the

mindful metacognition required for critical thinking. Therefore, joy is not a secondary outcome but a functional requirement for *deeper learning*.

Feasibility and the "basic topics" critique

One of the most persistent critiques of student-centred models is their perceived inability to handle "basic" topics like multiplication or initial shape recognition. While this study focused on Geometry, we argue that the mindful-meaningful-joyful cycle is equally applicable to arithmetic if the "Meaningful" component is grounded in grouping and repeated addition rather than static tables. Furthermore, our diagnostic assessment ensured that students had the necessary primitives, addressing the concern that higher-order models might leave struggling students behind. The extraordinary effect size ($d = 7.36$) suggests that when these prerequisites are met, the model provides a powerful accelerator for cognitive growth across all achievement levels.

Methodological reflection: Limitations and future directions

Despite the statistically significant findings and extraordinary effect size, several limitations must be addressed to contextualise the results. First, the one-group pre-test-post-test design lacks a comparison group, making it difficult to fully isolate the model's effects from threats to internal validity such as maturation, history, or testing effects. Second, the intensive Lesson Study format, involving seven collaborating teachers, likely induced a Hawthorne effect, where the high-focus environment influenced participant performance beyond the pedagogical model itself. Furthermore, the small sample size ($n = 20$) and the purposive selection of innovation-oriented teachers (Table 1) limit the generalizability of the findings to more traditional or resource-constrained settings.

Future research should therefore employ quasi-experimental designs with matched control groups to provide a more rigorous comparison of the model's effectiveness against traditional curricula. Longitudinal studies are also necessary to assess the sustainability of critical thinking gains over longer periods. Furthermore, researchers should explore the model's applicability across different mathematical domains (such as abstract algebra or number theory) and in varied cultural contexts to verify its universal transferability. Finally, investigating the model's implementation by a single teacher without the intensive support of a Lesson Study team would provide vital insights into its standalone feasibility in everyday classroom practice.

Transferability and the Indonesian context

The success of the model was undoubtedly aided by the Indonesian cultural concept of *gotong royong* (communal cooperation) (Murtadlo et al., 2024), which naturally aligns with the Lesson Study format. However, the core principles (affective safety, metacognitive awareness, and contextual anchoring) are universally transferable. For international educators, the implication is that solving the "critical thinking gap" requires a holistic shift that addresses the heart (affect) just as much as the head (cognition).

Conclusion

This study successfully designed and validated a pedagogical model that significantly enhances elementary students' critical thinking. By rejecting the false dichotomy between rigour and enjoyment, this framework shows that a holistic pedagogy—engaging the mind (mindful), connecting to the world (meaningful), and inspiring the heart (joyful)—is key to unlocking deep learning potential. The findings demonstrate that mathematics instruction can move from a focus on isolated procedural fluency to a synergistic focus on conceptual understanding. By creating an affective gateway through joy, students were able to employ mindful metacognition to solve meaningful real-world problems. Ultimately, this research provides an operational blueprint for educators to bridge the "knowing-doing gap" and cultivate the higher-order reasoning required in the 21st century.

Acknowledgments

We would like to thank the Directorate General of Research and Development of the Ministry of Higher Education, Science and Technology of the Republic of Indonesia for the 2025 Fundamental Research Grant Funding under contract numbers 014/LL17/DT.05.00/PL/2025 and 021/LPPM-LP/VI/2025.

Data Availability Statement

The article is based on data fully presented and discussed within the article itself; therefore, no additional data archiving is required.

References

Ang, C., Embi, M., and Md Yunus, M. (2016). Enhancing the quality of the findings of a longitudinal case study: Reviewing trustworthiness via ATLAS.ti. *The Qualitative Report, 21*(10), 1855–1867. https://doi.org/10.46743/2160-3715/2016.2480

Beattie, S. (2024). Play. In *Challenge Based Learning: Engaging with Students Through Interactivity* (pp. 135-175). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-96-0198-1_4

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

Chen, S. Y., Chen, W. C., and Lai, C. F. (2025). Generative AI as a reflective scaffold in a UAV-based STEM project: A mixed-methods study on students' higher-order thinking and cognitive transformation. *Education and Information Technologies*, *30*(17), 24787–24814. https://doi.org–/10.1007/s10639-025-13758-4

Clanton Harpine, E. (2024). Creating an intrinsically motivating learning environment: Promoting student engagement and intrinsic motivation. In *Service Learning in Higher Education: From pedagogy to practice* (pp. 59–76). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-51378-7_5

Cole, R. (2024). Inter-rater reliability methods in qualitative case study research. *Sociological Methods & Research*, *53*(4), 1944–1975. https://doi.org/10.1177/00491241231156971

Creswell, J. W., and Creswell, J. D. (2009). *Research Design: Qualitative, quantitative, and mixed methods approach* (3rd ed.). Sage Publications.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. https://doi.org/10.1007/BF02310555

Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher, 18*(3), 4–10. https://doi.org/10.3102/0013189X018003004

Fitrah, M., Caly Setiawan, Widihastuti, Marinding, Y., and Herianto. (2024). Evaluation of digital technology management in mathematics learning: A sequential explanatory design in eastern Indonesia. *Nordic Journal of Comparative and International Education (NJCIE), 8*(3). https://doi.or–g/10.7577/njcie.5926

Foster, C. (2018). Developing mathematical fluency: Comparing exercises and rich tasks. *Educational Studies in Mathematics, 97*(2), 121–141. https://doi.org/10.1080/0020739X.2013.770089

Fullan, M. (2016). *The New Meaning of Educational Change*. Teachers College Press.

Haapasalo, L., and Kadijevich, D. (2000). Two types of mathematical knowledge and their relation. *Journal für Mathematik-Didaktik*, *21*(2), 139–157. https://doi.org/10.1007/BF03338914

Jankowski, T., and Holas, P. (2014). Metacognitive model of mindfulness. *Consciousness and Cognition*, *28*, 64–80. https://doi.org/10.1016/j.concog.2014.06.005

Jonassen, D. H., and Strobel, J. (2006). Modeling for meaningful learning. In D. Hung and M. S. Khine (eds), *Engaged Learning with Emerging Technologies* (pp. 1–27). Dordrecht: Springer Netherlands. https://doi.org/10.1007/1-4020-3669-8

Kallio, H., Pietilä, A. M., Johnson, M., and Kangasniemi, M. (2016). Systematic methodological review: Developing a framework for a qualitative semi-structured interview guide. *Journal of Advanced Nursing, 72*(12), 2954–2965. https://doi.org/10.1111/jan.13031

Kannadass, P., Hidayat, R., Siregar, P. S., and Husain, A. P. (2023). Relationship between computational and critical thinking towards modelling competency among pre-service mathematics teachers. *TEM Journal, 12*(3), 1370–1382. https://doi.org/10.18421/TEM123-17

Kline, R. (2013). Exploratory and confirmatory factor analysis. In Y. Petscher, c. Schatschneider, and D. L. Compton (eds.), *Applied Quantitative Analysis in Education and the Social Sciences* (pp. 171–207). Routledge. https://doi.org/10.4324/9780203108550

Kudesia, R. S. (2019). Mindfulness as metacognitive practice. *Academy of Management Review*, *44*(2), 405–423. https://doi.org/10.5465/amr.2015.0333

Levin, M. (2018). Conceptual and procedural knowledge during strategy construction: A complex knowledge systems perspective. *Cognition and Instruction*, *36*(3), 247–278. https://doi.o–rg/10.1080/07370008.2018.1464003

Li, C., and Wen, F. (2023). Promoting deep learning in mathematics education —based on understanding by design theory. *Academic Journal of Mathematical Sciences, 4*(4), 57–61. https://doi.org/1–0.25236/AJMS.2023.040409

McEachen, J., Fullan, M., and Quinn, J. (2018). *NPDL Global Report. New Pedagogies for Deep Learning: A global partnership*. Deep Learning Series, 5. https://deep-learning.global/wp-content/uploads/2016/12/npdl-global-report-2016.pdf

Murtadlo, M., Albana, H., Helmy, M. I., Libriyanti, Y., Izazy, N. Q., and Saloom, G. (2024). Preserving the gotong royong character for Indonesian Gen-Z in the digital era. *International Journal of Evaluation and Research in Education*, *13*(3), 1631–1640. http://doi.org/10.11591/ijer–e.v13i3.27175

Mystakidis, S., Berki, E., and Valtanen, J. (2019). The patras blended strategy model for deep and meaningful learning in quality life-long distance education. *Electronic Journal of E-Learning, 17*(2), 66–78. https://doi.org/10.34190/JEL.17.2.01

Nguyen, P. N. T., Tran, K.-T., and Buu Le, G. (2025). Enhancing learning performance in primary education: The roles of problem solving and creative thinking challenges. *Journal of Elementary Education, 18*(3), 305–324. https://doi.org/10.18690/rei.4930

Polatbekova, K., Yergubekova, Z., Bostanova, A., Shektibayev, N., and Abdigapbarova, U. (2025). Development of the scientific and innovative potential of future teachers. *Scientific Reports*, *15*(1), 38665. https://doi.org/10.1038/s41598-025-22379-3

Richards, J. B., Hayes, M. M., and Schwartzstein, R. M. (2020). Teaching clinical reasoning and critical thinking: from cognitive theory to practical application. *Chest*, *158*(4), 1617–1628. https://doi.org/10.1016/j.chest.2020.05.525

Rittle-Johnson, B., and Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology*, *91*(1), 175. https://doi.org/–10.1037/0022-0663.91.1.175

Sablić, M., Mirosavljević, A., and Marinac, A. M. (2025). Students' perceptions of peer feedback in lower elementary school grades. *Journal of Elementary Education, 18*(2). https://doi.org/1–0.18690/rei.3773

Sarwanto, S., Fajari, L. E. W., and Chumdari, C. (2021). Critical thinking skills and their impacts on elementary school students. *Malaysian Journal of Learning and Instruction, 18*(2), 161–188. https://doi.org/10.32890/mjli2021.18.2.6

Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika, 52*(3–4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591

Stein, M. K., Grover, B. W., and Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal, 33*(2), 455–488. https://doi.org/10.3102/000283120–33002455

Thinwiangthong, S., Chatathicoon, S., Ya-amphan, D., and Somumchan, V. (2021). Early childhood students' mathematical ability in highscope and lesson study context. *Kasetsart Journal of Social Sciences*, *42*(3), 550–557. https://doi.org/10.34044/j.kjss.2021.42.3.15

Usiskin, Z. (1982). *Van Hiele Levels and Achievement in Secondary School Geometry*. University of Chicago. http://ucsmp.uchicago.edu/resources/van_hiele_levels.pdf

Van Den Heuvel-Panhuizen, M. (2003). The didactical use of models in realistic mathematics education: An example from a longitudinal trajectory on percentage. *Educational Studies in Mathematics*, *54* (1), 9-35. https://doi.org/10.1023/B:EDUC.0000005212.03219.dc

Vila-Henninger, L. A. (2019). Turning talk into "rationales": Using the extended case method for the coding and analysis of semi-structured interview data in ATLAS.ti. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique, 143*(1), 28–52. https://doi.org/10.1177/07–59106319852887

Wang, Q., and Abdullah, A. H. (2024). Enhancing students' critical thinking through mathematics in higher education: A systemic review. *SAGE Open*, *14*(3), 21582440241275651. https://doi.org/10.1177/21582440241275651
Willis, J. (2007). Cooperative learning is a brain turn-on. *Middle School Journal*, *38*(4), 4–13. https://doi.org/10.1080/00940771.2007.11461587

Appendices

Appendix A. Sample test items

Item 4 (Indicator: inference/deduction - Level: C5)

Context: You are designing a rectangular garden for a school project. You have exactly 24 meters of decorative fence.

1. Draw and label three different rectangular layouts that use the full 24m fence.
2. Calculate the area for each layout.
3. Which layout provides the maximum space for planting? Explain the relationship between the dimensions and the area you discovered.

Item 7 (Indicator: Evaluation/creation - Level: C6)

Context: A school playground consists of a large square area ($10m \times 10m$) and a small rectangular area ($4m \times 3m$) attached to one side.

1. Create a scale drawing of this composite playground.
2. A contractor suggests that doubling the length of the small rectangle will double the total area of the playground. Evaluate this claim using mathematical proof.
3. Propose a new dimension for the small rectangle that would increase the total playground area by exactly 25%.

Appendix B. Critical thinking scoring rubric

Table 7

Critical thinking scoring rubric (Scale 1-4)

Score	Performance level	Description of mathematical reasoning
4	Exemplary	Correct mathematical solution. Provides a robust, logical justification connecting procedural steps to conceptual relations. Demonstrates clear inference/creation.
3	Proficient	Correct mathematical solution. Provides a partial justification or verbalisation of the process but may lack depth in conceptual connection.
2	Developing	Partially correct solution. Reasoning is mostly procedural or based on rote recall of formulas without clear contextual justification.
1	Beginning	Incorrect solution or purely algorithmic attempt with no justification. Shows high dependence on memorized facts without understanding.

Appendix C. Instructional observation matrix

Table 8

Instructional observation matrix

Principle	Score 1 (*Transmissionist*)	Score 4 (*Deeper learning*)
Mindful	Teacher provides all steps. Students follow instructions passively.	Teacher uses "What if?" prompts. Students actively monitor and justify their own strategy selection.

Meaningful	Math is taught as abstract symbols. Problems are purely numerical.	Math is anchored in tangible classroom/real-world contexts. Students use concrete media to bridge to abstraction.
Joyful	High-stakes environment. Fear of being "wrong." Low student energy.	Psychologically safe 'blame-free' zone. Students engage as 'detectives' or 'designers.' High collaborative energy.

Authors:

Pariang Sonang Siregar, PhD
Associate Professor, Faculty of Educational Sciences, Universitas Rokania, Raya Pasir Pengaraian Street, KM 15, Rokan Hulu, Republic of Indonesia, E-mail: pariangsonangsiregar@gmail.com
izredni profesor, Fakulteta za edukacijske vede, Univerza Rokania, Raya Pasir Pengaraian Street, KM 15, Rokan Hulu, Republika Indonezija, E-pošta: pariangsonangsiregar@gmail.com

Neni Hermita, PhD
Professor, Faculty of Teacher Training and Education, Universitas Riau, Kampus Bina Widya, KM 12, Pekanbaru 28293, Republic of Indonesia, E-mail: neni.hermita@lecturer.unri.ac.id
redna profesorica, Fakulteta za izobraževanje učiteljev in pedagogiko, Univerza Riau, Kampus Bina Widya, KM 12, 28293 Pekanbaru, Republika Indonezija, E-pošta: neni.hermita@lecturer.unri.ac.id

Firman Santosa
Lecturer, Faculty of Computer Science, Universitas Rokania, Raya Pasir Pengaraian Street, KM 15 Rokan Hulu, Republic of Indonesia, E-mail: firman.snts@gmail.com
Predavatelj, Fakulteta za računalništvo, Univerza Rokania, Raya Pasir Pengaraian Street, KM 15, Rokan Hulu, Republika Indonezija, E-pošta: firman.snts@gmail.com

Tety Herti
Student, Faculty of Educational Sciences, Universitas Rokania, Raya Pasir Pengaraian Street, KM 15, Rokan Hulu, Republic of Indonesia, E-mail: tetyherti2873@gmail.com
Študentka, Fakulteta za edukacijske vede, Univerza Rokania, Raya Pasir Pengaraian Street, KM 15, Rokan Hulu, Republika Indonezija, E-pošta: tetyherti2873@gmail.com