Journal of Honai Math, 8(3), 481-500, December 2025
https:/doi.org/10.30862/jhm.v8i3.1014

Bridging numeracy and critical reasoning through realistic mathematics education: evidence from a mixed-methods classroom

Rustam Effendy Simamora *, Heldina L.Toruan

Mathematics Education Department, Universitas Borneo Tarakan, Tarakan, Indonesia

* Correspondence: erustam@borneo.ac.id

Received: 15 August 2025 | Revised: 4 November 2025 | Accepted: 30 December 2025 | Published: 30 December 2025
© The Author(s) 2025

Abstract

Critical thinking in numeracy contexts (CTNC) is a core twenty-first-century competency, yet its systematic development remains limited in conventional mathematics instruction. Realistic Mathematics Education (RME) offers a theoretically grounded approach to fostering CTNC through contextualised problem solving, guided reinvention, and collaborative discourse. This mixed-methods study investigated the effectiveness of RME in enhancing CTNC among junior secondary students and examined their experiential learning processes. Thirty-eight Year 7 students (n = 19 per class) were selected via cluster sampling. A quasi-experimental, non-randomised control group design with pre-test and post-test measures was employed in the quantitative phase. The qualitative phase involved three waves of semi-structured interviews with six purposively selected students, using member checking and theoretical sampling until data saturation was achieved. Results indicated a statistically significant improvement in CTNC for students in the RME group compared to those receiving conventional instruction, $t(36) = 5.55$, $p < .001$. Thematic analysis revealed three interconnected dimensions: enhanced contextual sense-making through engagement with realistic tasks; deepened critical reasoning facilitated by structured peer discourse; and increased metacognitive awareness during strategy evaluation. Students also reported greater engagement and perceived relevance of mathematics to everyday contexts. These findings suggest that RME effectively strengthens CTNC by linking abstract numeracy to authentic experiences through collaborative inquiry. The study highlights the importance of systematically integrating context-rich, discussion-oriented pedagogies into secondary mathematics curricula.

Keywords: Contextualised problem solving; Critical thinking in numeracy contexts; Realistic mathematics education; Mathematical discourse; Mixed-methods research

Introduction

Critical Thinking in Numeracy Contexts (CTNC) has increasingly been recognised as a foundational competency for the twenty-first century. CTNC refers to the capacity, disposition, and skill to apply mathematical reasoning in solving authentic everyday problems (Geiger et al., 2015). Grounded in a critical orientation, CTNC extends beyond procedural computation to encompass the analytical interpretation and evaluative use of numerical information for informed decision-making. Substantial evidence indicates that strong critical thinking competencies are associated with improved outcomes across multiple domains, including health, finance, civic participation, and educational attainment (Butler et al., 2017; Gelerstein et al., 2016). Nevertheless, despite its recognised significance, international assessments continue to report relatively low levels of numeracy and critical reasoning among Indonesian students (Wijaya et al., 2024). This concern is further reinforced by preliminary classroom-based investigations conducted in regions such as Tarakan, North Kalimantan, which similarly identified limited student capacity in numeracy-related critical reasoning (Luthfiyah et al., 2024; Rizmawati et al., 2024). These conditions underscore the urgent need for pedagogical approaches that systematically integrate mathematical understanding with authentic problem solving and critical reflection in order to cultivate CTNC effectively.

Within mathematics education, CTNC may be understood as the intersection between numeracy and critical thinking. It requires students not only to interpret contextual information but also to evaluate the relevance of available data, construct logically justified solutions, and generate alternative strategies when appropriate (Jain & Rogers, 2019; Watson, 2002). In the present study, CTNC is conceptualised through five interrelated cognitive indicators: (1) summarising essential information embedded within numeracy problems; (2) connecting relevant information to problem objectives while selecting appropriate solution strategies; (3) formulating solutions supported by logical reasoning; (4) evaluating the adequacy and relevance of available information; and (5) proposing more effective and efficient alternative solutions. The development of these competencies necessitates learning environments that extend beyond repetitive procedural practice by engaging students in authentic mathematical modelling, reflective discourse, and strategic evaluation.

Realistic Mathematics Education (RME) provides a theoretically robust instructional framework that aligns closely with these educational demands. Originating from the Freudenthal Institute, RME emphasises the use of “realistic” contexts namely, situations that are experientially meaningful and cognitively imaginable for students as the starting point for mathematical exploration (Van den Heuvel-Panhuizen, 2020). Through the process of guided reinvention, students progressively transform informal experiences into formal mathematical understanding, supported by the dynamic interaction between models of contextual representations and models for abstract mathematical reasoning (Heuvel-Panhuizen, 2003). Such instructional characteristics inherently activate the core dimensions of CTNC. Contextual problem solving requires students to decode information and determine suitable strategies; collaborative discussion and comparison of solutions promote logical reasoning and evaluative thinking; and the transition from informal modelling to formal abstraction encourages students

to generate increasingly efficient and sophisticated alternative solutions. In this sense, RME functions not merely as a content-delivery approach, but as a cognitive scaffold that positions numeracy as a tool for critical inquiry and responsible decision-making.

Despite the strong theoretical correspondence between RME and CTNC, empirical investigations explicitly examining the influence of RME on CTNC remain notably limited. As of mid-2025, systematic literature searches revealed an absence of studies positioning CTNC as the primary outcome variable within RME-based interventions. Existing mixed-methods studies have generally focused on broader constructs such as critical thinking or mathematical problem solving, yet many have not incorporated rigorous qualitative validation procedures, including member checking and theoretical sampling, to substantiate the instructional mechanisms underlying cognitive development (Charmaz, 2014; Creswell & Creswell, 2023). This gap highlights the need for a methodologically rigorous design capable of both quantifying CTNC development and capturing students' experiential learning trajectories in depth. Accordingly, the present study employs an embedded experimental mixed-methods design. Quantitative pre-test and post-test data are utilised to examine the magnitude of CTNC improvement, while qualitative interviews are conducted to illuminate the cognitive and affective processes through which such development occurs. Finally, guided by this framework, the study addresses two principal research questions:

1. Is there a significant difference in CTNC improvement between students taught through the RME approach and those receiving conventional mathematics instruction?
2. How do students perceive and experience RME instruction, and in what ways do these experiences contribute to the development of CTNC?

Methods

This study employed a concurrent mixed-methods design with an embedded quantitative component (Creswell, 2023), in which quantitative data served as the primary source of evidence and qualitative data functioned to provide complementary explanatory insights. The quantitative strand adopted a quasi-experimental, non-randomised control group pre-test–post-test design (Ary et al., 2019) to examine the effectiveness of RME in enhancing students' CTNC. The qualitative strand explored students' perceptions and experiences of RME instruction, particularly regarding the processes through which CTNC developed during learning activities.

The study was conducted in two phases. The first phase, spanning May 2023 to February 2024, involved the development and validation of research instruments and instructional materials. The second phase consisted of the classroom intervention implemented in March 2024 over six sessions, including one pre-test session, four instructional sessions, and one post-test session.

Participants and Sampling

The target population comprised all Year 7 students (N = 112) enrolled at the participating secondary school. Two intact classes were selected through cluster sampling procedures. The experimental group (n = 27) received instruction based on the RME approach, whereas the control group (n = 26) received conventional mathematics instruction. Following data screening for attendance consistency and assessment completeness, the final analytical sample consisted of 38 students, with 19 students in each group.

For the qualitative strand, six students from the experimental group were selected purposively using theoretical sampling procedures across two recruitment waves. Participant selection was stratified according to changes in CTNC scores in order to capture diverse learning trajectories and variations in instructional experiences (see Table 1). Each recruitment wave included one participant demonstrating high score gains, one participant exhibiting low or stagnant gains, and one participant showing minimal or negative gains.

Table 1. Demographic Characteristics of Interview Participants

Recruitment Wave	Participant Code	CTNC Score Change (points)	Gender	Age (years)
Wave 1	PQ#1	+37	Female	12
	PQ#2	+2	Female	13
	PQ#3	-3	Female	13
Wave 2	PQ#4	+34	Female	13
	PQ#5	+10	Female	14
	PQ#6	+8	Male	13

Instruments and Validation

CTNC was assessed using a researcher-developed test instrument contextualised within financial mathematics topics, including buying and selling, profit and loss, discounts, and taxation. The instrument was explicitly aligned with five CTNC indicators: (1) identifying relevant information; (2) linking information to analytical objectives; (3) formulating appropriate solution strategies; (4) evaluating the adequacy and relevance of information; and (5) proposing effective and efficient alternative solutions. A sample post-test item is presented in Figure 1

Pak Budi is a fruit seller. One day he bought ten baskets of oranges for a total of Rp300,000.00. Each basket contains 15 kg of fruit. He paid Rp150,000.00 for transportation. To make a profit, Pak Budi's wife suggests two selling-price options:

(a) sell the oranges at Rp3,500.00 per kilogram, or
(b) sell all the oranges together for a total of Rp525,000.00.

However, according to Pak Budi's calculations, he would actually incur a loss with those prices. Determine whether Pak Budi will indeed suffer a loss by answering the following questions:

a. State the important information given in the problem.
b. How would you determine whether Pak Budi incurs a loss based on the given information? Explain your strategy or method.
c. Are the two suggestions from Pak Budi's wife (items a and b above) overlapping statements? Explain your reasoning.
d. Carry out the calculations following your strategy from item b, and conclude whether Pak Budi does incur a loss. Include your explanation.
e. Provide an alternative method to determine whether Pak Budi incurs a loss, then state which method is more time-efficient: the method in item d or your new method.

Figure 1. Example of a Post-test Item

Content validity was established through expert judgement involving two mathematics education specialists and one experienced junior secondary mathematics teacher, resulting in a mean validity score of 3.74 out of 4.00. Item analysis indicated moderate difficulty levels and moderate-to-excellent discrimination indices across test items. Reliability analysis demonstrated high internal consistency, with Cronbach's α coefficients of .913 for the pre-test and .927 for the post-test, both categorised as "very high" according to George and Mallery.

Learning Materials and Intervention

The instructional materials consisted of RME-based and conventional lesson plans (LPs), student worksheets (SWs), and a lesson observation implementation protocol (LOIP). Both instructional approaches addressed the same financial mathematics topics across four 60-minute sessions.

The RME lesson plans were designed according to Gravemeijer's (1994) instructional sequence, comprising: (1) contextual problem introduction; (2) collaborative group exploration; (3) guided teacher facilitation; (4) intergroup discussion and comparison of strategies; and (5) formalisation and consolidation of mathematical concepts. In contrast, the conventional lesson plans followed a direct-instruction format consisting of teacher explanation, worked examples, and individual practice activities. Furthermore, validation results demonstrated high levels of instructional feasibility across all materials, including the RME lesson plans (M = 3.82), conventional lesson plans (M = 3.79), student worksheets (M = 3.80), and LOIP instrument (M = 3.92).

The student worksheets (see Figure 2) contained eight realistic and contextually grounded mathematical tasks distributed across four instructional meetings. These activities were specifically designed to support horizontal mathematisation processes and facilitate the

development of CTNC competencies. The LOIP instrument was employed to monitor implementation fidelity and ensure adherence to the planned instructional phases.

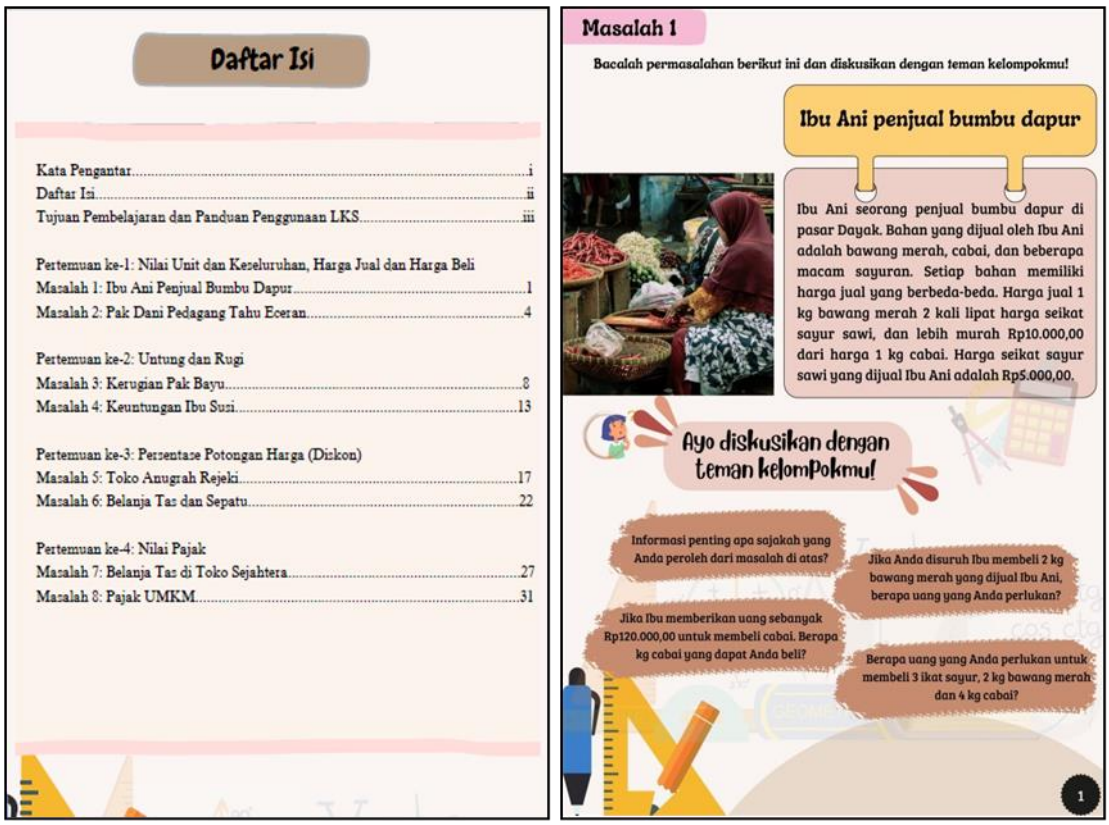


Figure 2. Overview of the Table of Contents and Sample Tasks in the RME-Based Student Worksheets

Data Collection and Analysis Procedures

Quantitative data were collected through pre-test and post-test administration in both groups. Descriptive statistical analyses were first conducted to characterise baseline and post-intervention performance. Prior to inferential analysis, statistical assumptions were examined using the Shapiro–Wilk test for normality and Levene's test for homogeneity of variance. Gain scores were subsequently calculated and analysed using an independent samples t-test.

The null hypothesis (H_0) stated that no significant difference would exist in CTNC gain scores between students receiving RME instruction and those receiving conventional instruction. Conversely, the alternative hypothesis (H_1) proposed that students in the RME group would demonstrate significantly greater improvements in CTNC.

Qualitative data were obtained through three sequential stages of semi-structured interviews. The first stage explored students' prior experiences in mathematics learning, perceptions of RME instruction, levels of engagement, and perceived challenges. The second stage involved member checking interviews with the same participants to verify the credibility and accuracy of preliminary interpretations. The third stage employed theoretical sampling with three additional participants to evaluate data saturation (Charmaz, 2014).

All interviews were conducted face-to-face in Indonesian and lasted approximately 20–30 minutes. The interviews were audio-recorded, transcribed verbatim, and analysed using thematic analysis procedures following Nowell et al. (2017), with NVivo® software supporting data management and coding processes. Analysis progressed iteratively from open coding to the development of subthemes and overarching thematic categories.

Research Legitimation

Methodological rigour was established through the integration and triangulation of quantitative and qualitative procedures. The quality of the quantitative instruments was ensured through expert validation, item analysis, and reliability testing. The trustworthiness of the qualitative findings was strengthened through member checking, time triangulation, audit trails, and reflexive documentation of potential researcher bias (Creswell, 2018). Collectively, these legitimation strategies enhanced the credibility, consistency, and accountability of interpretations derived from both quantitative and qualitative data sources.

Results

This section presents the findings from both the quantitative and qualitative phases of the study. The results are organised into three sections: (1) implementation of the RME intervention; (2) quantitative findings concerning the improvement of CTNC; and (3) qualitative findings related to students' learning experiences during RME instruction.

Implementation of the RME intervention

The RME intervention was implemented across four 60-minute instructional sessions covering the following topics: (1) buying and selling, including profit and loss; (2) percentage calculations; (3) discounts; and (4) the integration of tax and discount calculations. Each instructional session followed the core RME sequence consisting of contextual problem introduction, individual exploration, collaborative group discussion, intergroup comparison, and formal mathematical consolidation.

Among the four sessions, the third meeting, which focused on percentage discounts, demonstrated the highest implementation fidelity score (3.80 out of 4.00). This session incorporated two contextualised tasks: Problem 5: Anugrah Rejeki Store and Problem 6: Shopping for Bags and Shoes. Figure 3 presents Problem 5 (Task 1), in which students were required to calculate discount values and compare final prices across different promotional scenarios within a retail shopping context.

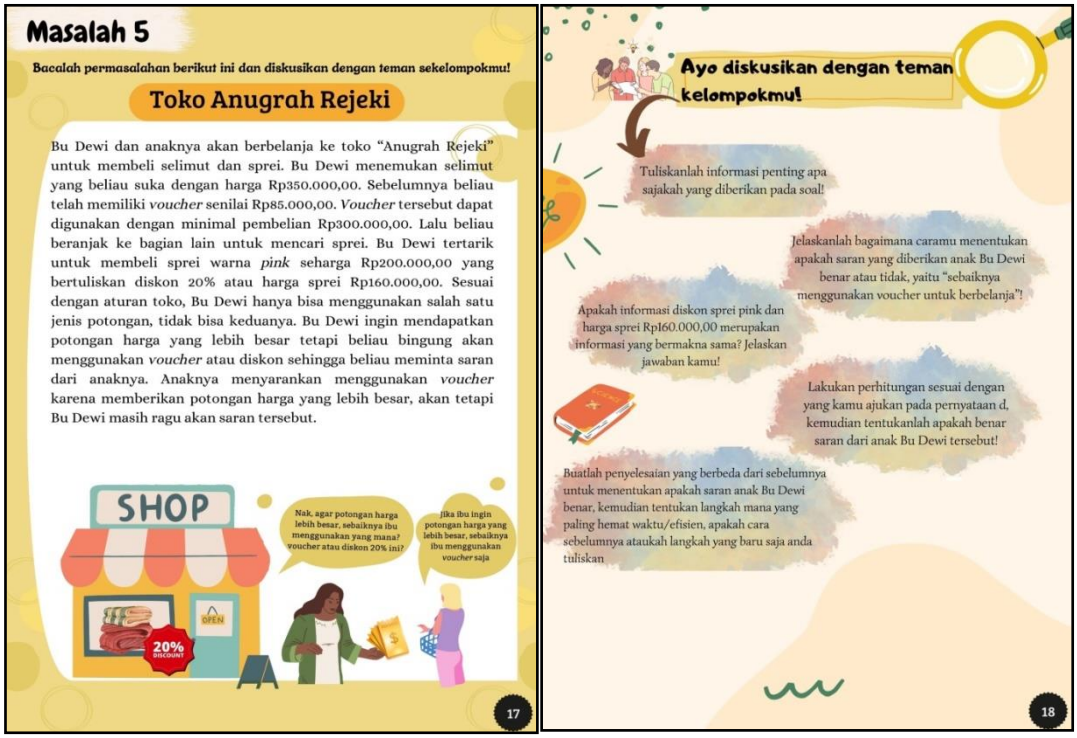


Figure 3. Problem 5 (Task 1): Discount Calculation at Anugrah Rejeki Store

The task situated students within an authentic shopping scenario requiring them to: (a) identify relevant information regarding original prices and discount percentages; (b) calculate discount values accurately; (c) determine final prices following discount application; and (d) compare the relative advantages of alternative promotional offers. Consequently, the task engaged multiple CTNC indicators, particularly those associated with evaluating information relevance, selecting appropriate calculation strategies, and proposing efficient solution procedures.

During implementation, students worked collaboratively in groups of four to five members. The teacher facilitated conceptual discussion through guiding questions such as, "How can you determine which discount offers the greatest benefit?" and "Are there alternative ways to calculate the final price?" Students were first encouraged to think individually before engaging in collaborative problem solving. Group representatives subsequently presented their reasoning to the class, after which whole-class discussions were conducted to compare strategies and evaluate solutions. The instructional session concluded with teacher-guided formalisation of the discount formula:

$$Final\ Price = Original\ Price \times (1 - Discount\ Rate)$$

Quantitative Results

Descriptive analyses were conducted to examine changes in students' CTNC performance before and after the instructional intervention in both groups. The experimental group received RME-based instruction, whereas the control group participated in conventional learning (CL). The experimental group obtained a mean pre-test score of 14.79 (SD = 8.22), while the control

group demonstrated a higher initial mean score of 20.47 (SD = 6.98). Following the intervention, however, the experimental group exhibited a substantial increase in performance, achieving a mean post-test score of 37.32 (SD = 9.09), compared with 25.89 (SD = 6.92) in the control group. Analysis of gain scores further demonstrated the magnitude of improvement experienced by each group. The experimental group achieved an average gain of 22.53 points, whereas the control group improved by only 5.42 points. These findings suggest that the RME approach was considerably more effective than conventional instruction in enhancing students' CTNC.

Furthermore, prior to inferential analysis, assumptions underlying parametric testing were examined. The Shapiro–Wilk normality test indicated that gain scores in both groups were normally distributed ($p > .050$). In addition, Levene's test demonstrated homogeneity of variance across groups ($p = .489$). These results confirmed that the data satisfied the assumptions required for an independent samples t-test. The independent samples t-test revealed a statistically significant difference in CTNC gain scores between the experimental and control groups, $t(36) = 5.55$, $p < .001$. The observed mean difference of 17.11 points indicates that students receiving RME instruction demonstrated substantially greater improvement in CTNC than those receiving conventional instruction. These findings suggest that RME effectively facilitates the development of multiple CTNC dimensions, including the capacity to analyse numerical information, select relevant problem-solving strategies, construct logically justified mathematical arguments, and evaluate solutions within meaningful contexts. Importantly, the effectiveness of the intervention was reflected not only in higher post-test achievement but also in markedly greater learning gains, despite the experimental group beginning with lower baseline scores. Overall, the quantitative findings reinforce the proposition that context-based and student-centred instructional approaches such as RME can substantially enhance students' CTNC.

Quantitative Results

The qualitative analysis aimed to explore students' perceptions of RME instruction and its contribution to the development of CTNC. Semi-structured interviews were conducted with six students from the experimental group across two interview stages following the fourth and sixth instructional meetings. Thematic analysis generated six overarching themes: (1) perceptions of mathematics and mathematics learning; (2) impressions of RME instruction; (3) perceived development of CTNC; (4) engagement during learning activities; (5) responses to realistic mathematical problems; and (6) challenges encountered during the learning process.

Perceptions of Mathematics and Mathematics Learning

Most participants expressed positive attitudes towards mathematics despite perceiving it as a challenging subject. For example, PQ#1 stated:

> *"I enjoy learning mathematics. Even though it is very difficult, I still like the challenge. When solving problems, I feel satisfied when I find the answer" (PQ#1, Intv. 1).*

Similarly, PQ#4 explained:

> *"I like mathematics, but not when the answers are difficult. I like it when I can understand" (PQ#4, Intv. 3).*

At the same time, most participants described their prior learning experiences as predominantly conventional and teacher-centred. PQ#4 noted:

> *"So far, we have learned mathematics through lectures. The teacher explains the material, and then we work on practice questions" (PQ#4, Intv. 3).*

These responses suggest that students' previous learning experiences provided limited opportunities for interaction, discussion, and reflective reasoning, thereby constraining the development of CTNC.

Impressions of Learning Through the RME Approach

Most students responded positively to the RME approach. PQ#2 described RME instruction as enjoyable because of its collaborative nature:

> *"I like RME more because we learn in groups. It is enjoyable to find the answers together" (PQ#2, Intv. 2).*

Likewise, PQ#1 emphasised the supportive atmosphere created through discussion activities:

> *"I prefer group discussions because working together feels lively. It does not feel too difficult or burdensome" (PQ#1, Intv. 1).*

Nevertheless, not all participants preferred collaborative learning environments. PQ#3 expressed discomfort with group-based interaction:

> *"I prefer studying alone because I do not want others to see my answers, especially since my group members were joking around yesterday" (PQ#3, Intv. 2).*

This finding indicates that the effectiveness of RME implementation is influenced not only by instructional design but also by classroom social dynamics and students' individual preferences regarding interaction and participation.

Perceived Development of CTNC

All participants reported experiencing some degree of improvement in their critical thinking abilities, particularly in analysing information and constructing arguments. PQ#1 explained:

> *"I think I improved in giving arguments because the teacher trained us to do that. I still find it difficult, but I feel that I have improved" (PQ#1, Intv. 1).*

However, improvement was not equally evident across all CTNC dimensions. PQ#3, for instance, acknowledged difficulty in generating alternative strategies:

> *"It is still just average. I cannot determine the right strategy" (PQ#3, Intv. 1).*

Several students also reported difficulty verbalising their reasoning processes despite understanding the calculations conceptually. PQ#1 stated:

> *"I am not good at explaining things. Even when I understand the calculations, I cannot easily express them in words" (PQ#1, Intv. 1).*

These findings suggest that while RME supported aspects of critical reasoning development, improvements in mathematical communication skills were not uniformly experienced across participants.

Engagement During Learning Activities

The RME approach appeared to enhance students' active participation during instruction. Five of the six participants identified collaborative discussion as the most engaging component of the learning process. These responses indicate that collaborative strategies within RME can encourage participation, mutual support, and shared problem solving.

However, high engagement did not necessarily guarantee effective collaboration. PQ#3 criticised the group dynamics, stating:

> *"Yesterday, my group members were only playing around" (PQ#3, Intv. 2).*

This finding demonstrates that the effectiveness of collaborative learning depends substantially on group commitment and highlights the important role of the teacher in facilitating productive group interaction.

Responses to Realistic Problems

The realistic contexts embedded within the RME tasks were positively received by most students. PQ#1 explained:

> *"The realistic problems are interesting because they relate to everyday life, so the concepts are easier to understand" (PQ#1, Intv. 2).*

This response was particularly evident during the third instructional meeting, in which students engaged with discount-related problems that closely reflected their everyday shopping experiences.

Nevertheless, some participants experienced difficulty interpreting lengthy contextual narratives. PQ#2 stated:

> *"The realistic problems are okay, but sometimes I do not understand them" (PQ#2, Intv. 1).*

These findings suggest that although realistic contexts can support conceptual understanding, contextual problems must be carefully designed to remain cognitively accessible for students with varying levels of comprehension.

Difficulties Encountered During Learning

In addition to difficulties related to interpreting contextual problems and articulating reasoning, several students experienced challenges with fundamental arithmetic operations, particularly division. Weaknesses in basic computational fluency appeared to hinder students' ability to engage fully in higher-order reasoning processes, as cognitive attention was frequently directed towards calculation procedures rather than analytical evaluation.

PQ#3 reflected:

> *"When Ibu KZ taught us previously, she gave us the problems and then immediately provided the answers. We were not accustomed to calculating things ourselves" (PQ#3, Intv. 1).*

Similarly, PQ#4 stated:

> *"I do not really understand division because we were never properly taught it in middle school" (PQ#4, Intv. 3).*

These findings indicate that limitations in foundational computational skills may substantially reduce the effectiveness of RME-based learning. When students struggle with basic arithmetic operations, cognitive resources are diverted towards procedural calculation rather than critical reasoning, strategic evaluation, and reflective problem solving.

Discussion

The findings of this study provide strong empirical evidence that the RME approach significantly enhances students' CTNC. The substantial mean gain difference of 17.11 points, together with the statistically significant t-test result, $t(36) = 5.55$, $p < .001$, indicates that RME exerts a markedly stronger influence on CTNC development than conventional instructional approaches, even when students begin with lower baseline performance levels. Improvement was observed across all five CTNC indicators, namely information summarisation, strategy selection, reasoned solution formulation, information evaluation, and the generation of alternative solutions.

These quantitative findings are consistent with the conceptualisation of CTNC as a critical orientation towards numeracy proposed by Geiger et al. (2015), in which contextualised mathematical activity engages students in analytical and evaluative reasoning rather than procedural recall alone. Notably, the greatest improvements were evident in strategy formulation and logical reasoning, whereas progress related to evaluative dimensions was comparatively more modest. This pattern suggests that although RME naturally supports divergent reasoning and comparative strategy analysis through collaborative interaction, the development of meta-evaluative thinking may require more explicit instructional scaffolding. Such findings align with Facione et al.'s (2015) argument that structured argumentation and evaluative dialogue are essential for fostering critical judgement.

From a pedagogical perspective, the effectiveness of RME appears to stem from its capacity to situate mathematical reasoning within meaningful contexts while simultaneously encouraging reflective and dialogic engagement. The iterative movement between contextual interpretation, informal modelling, and formal mathematical abstraction enabled students to connect numerical procedures with analytical decision-making processes. In this sense, CTNC development occurred not merely through exposure to realistic problems, but through sustained participation in collaborative reasoning practices that required students to justify, compare, and refine their mathematical thinking.

Qualitative Insights: Engagement, Cognitive Friction, and Metacognitive Growth

The qualitative findings provide explanatory depth to the quantitative results by illuminating how students experienced and internalised RME instruction. Most participants reported heightened engagement during learning activities, particularly because the realistic scenarios rendered mathematical problems more meaningful and accessible. As PQ#5 remarked, "The problems are like real daily events." This perceived authenticity reflects the principle of horizontal mathematisation within RME, whereby students connect everyday experiences with formal mathematical reasoning processes (Van den Heuvel-Panhuizen & Drijvers, 2020).

Collaborative discourse also emerged as a significant mechanism supporting CTNC development. Students described how group discussions enabled them to exchange

perspectives, compare solution strategies, and construct shared understanding. PQ#2 explained, "We can help each other and understand how others think." Such interaction appears to support not only the cognitive dimensions of CTNC but also its dispositional aspects, including openness to alternative viewpoints and willingness to justify reasoning (Schoenfeld, 2020).

Nevertheless, the qualitative data also revealed several forms of cognitive friction that constrained CTNC development. Four participants reported difficulties with foundational arithmetic operations, particularly division, which diverted cognitive attention from higher-order analytical processes towards procedural computation. These findings help explain the relatively low pre-test performance observed among several students despite prior exposure to mathematics instruction and support Simamora et al.'s (2024) argument that deficiencies in basic numeracy can limit students' ability to engage critically with context-rich tasks.

In addition, one participant expressed a preference for conventional instruction due to discomfort with collaborative learning dynamics, suggesting that the effectiveness of RME is influenced by students' social and emotional readiness for interactive learning environments. This finding highlights the importance of fostering psychological safety, equitable participation, and productive group norms within collaborative classrooms. Furthermore, several students experienced difficulty articulating their reasoning verbally, indicating that the development of CTNC also depends upon explicit support for mathematical communication and academic language practices (Hammond & Gibbons, 2005).

Despite these challenges, the iterative instructional cycle characteristic of RME, engaging with realistic problems, discussing strategies, receiving feedback, and revising solutions, appeared to promote metacognitive awareness indirectly. The requirement to justify and compare strategies encouraged students to reflect critically upon their own reasoning processes, thereby fostering habits of self-monitoring and evaluative thinking that are foundational to sustained critical numeracy development (Plomp, 2013)

Theoretical and Pedagogical Implications

The convergence of quantitative effectiveness and qualitative experiential evidence positions RME not merely as an instructional strategy, but as a pedagogical framework capable of operationalising CTNC through guided reinvention, contextual modelling, and structured mathematical discourse. Theoretically, the study extends Geiger et al.'s (2015) model of critical numeracy by demonstrating how RME simultaneously activates both the cognitive tools and dispositional dimensions of numeracy within authentic learning contexts.

Several pedagogical implications emerge from these findings. First, instructional materials should embed explicit CTNC-oriented prompts within realistic mathematical tasks rather than assuming that contextualisation alone will stimulate critical reasoning. Well-designed learning materials should progressively require students to compare strategies, justify decisions, evaluate information, and consider alternative solutions.

Second, teacher professional development should move beyond content transmission towards discourse-oriented pedagogical facilitation. Teachers require sustained support in managing collaborative interaction, posing probing questions, and providing adaptive

scaffolding that balances cognitive challenge with accessibility (Vanden, 2020). In this regard, the teacher's role shifts from knowledge transmitter to facilitator of collective mathematical inquiry.

Third, the successful implementation of RME requires attention to students' foundational computational fluency. Without adequate arithmetic proficiency, students may experience cognitive overload that limits their capacity to engage in higher-order reasoning processes (Simamora et al., 2024). Consequently, diagnostic assessment and targeted remediation of basic numeracy skills should either precede or accompany RME implementation. Finally, the findings suggest the importance of cultivating classroom cultures that normalise iterative reasoning, reflective dialogue, and metacognitive questioning. Structured reflective routines, such as "I see, I think, I wonder," may help transform CTNC from an isolated performance outcome into an enduring intellectual habit (Wastom, 2024).

Limitations and Directions for Future Research

Despite its contributions, this study possesses several limitations that should be acknowledged. First, although the quasi-experimental design was appropriate within the constraints of an authentic classroom setting, the absence of random assignment limits the extent to which causal conclusions may be drawn, as uncontrolled external variables may have influenced the findings.

Second, the intervention was limited to financial mathematics topics and was implemented over a relatively short instructional duration. Consequently, the study may not fully capture the long-term retention or transferability of CTNC across broader mathematical domains. Future investigations should therefore examine the application of RME across additional areas such as geometry, algebra, and statistics, while also incorporating longitudinal designs to evaluate the sustainability of CTNC development over time.

Third, the qualitative component involved only six participants from the experimental group, thereby limiting the breadth of experiential variation represented within the findings and excluding perspectives from students receiving conventional instruction. Expanding participant diversity across socioeconomic backgrounds, achievement levels, and instructional contexts would provide richer insight into how classroom culture, institutional support, and adaptive task design mediate the effectiveness of RME.

Future research may also benefit from employing more rigorous experimental frameworks, such as randomised controlled trials or Solomon four-group designs, in order to strengthen internal validity (Creswell, 2023). In addition, incorporating teacher perspectives and classroom observational analyses may further illuminate the pedagogical mechanisms through which RME facilitates CTNC development. Finally, these directions would contribute towards the development of more responsive, scalable, and equitable approaches for integrating critical numeracy development within secondary mathematics education.

Conclusion

This mixed-methods study demonstrates that RME significantly enhances students' CTNC at the junior secondary level. Quantitative findings revealed statistically significant and substantively meaningful differences in CTNC improvement between students receiving RME instruction and those taught through conventional approaches, $t(36) = 5.55$, $p < .001$, with a mean gain difference of 17.11 points. Notably, students in the experimental group progressed from a lower pre-test baseline to substantially higher post-test achievement. These quantitative outcomes were reinforced by qualitative evidence indicating that students perceived RME as more engaging, meaningful, and supportive of analytical reasoning, particularly through the use of realistic problem contexts and collaborative mathematical discussion.

The integration of quantitative and qualitative findings highlights three principal mechanisms through which RME facilitates CTNC development. First, contextualised mathematical problems activate students' prior knowledge and encourage analytical interpretation of information situated within authentic scenarios. Second, guided reinvention and collaborative discussion processes scaffold students' abilities to select strategies, construct logical arguments, and compare alternative solutions. Third, the iterative cycle of exploration, verification, discussion, and formalisation promotes the development of metacognitive and evaluative habits that are central to critical numeracy practices.

At the same time, the qualitative findings identified several factors that moderated the effectiveness of the intervention. Deficiencies in foundational arithmetic skills, variability in collaborative group dynamics, and difficulties in verbalising mathematical reasoning constrained the extent to which some students were able to achieve optimal CTNC development. These findings suggest that the successful implementation of RME requires not only contextually rich instructional design but also adequate support for computational fluency, mathematical communication, and productive classroom interaction.

Several implications for mathematics education emerge from this study. First, mathematics instruction should incorporate core RME principles including contextual problem design, collaborative inquiry, and reflective formalization while simultaneously providing explicit scaffolding for communication and foundational numeracy skills. Second, curriculum developers and teacher educators should prioritise professional development initiatives that prepare teachers to facilitate mathematical discourse, manage collaborative learning environments, and design realistic tasks that balance cognitive challenge with accessibility. Third, the findings underscore the importance of cultivating classroom cultures that normalise reflective reasoning, strategic comparison, and critical evaluation as routine mathematical practices rather than isolated instructional activities.

Despite these contributions, several limitations should be acknowledged. The study focused exclusively on financial mathematics topics among Year 7 students, representing only a limited portion of the broader mathematics curriculum. Consequently, the transferability of the findings to other mathematical domains, such as algebra, geometry, or statistics, remains uncertain. In addition, the intervention was implemented over a relatively short duration, which may not have allowed sufficient time for students to internalise the instructional approach fully

or demonstrate sustained long-term CTNC development. The cultivation of critical thinking habits requires consistent and prolonged engagement across mathematical topics and learning experiences.

Furthermore, the study was conducted within a single public secondary school in Tarakan using a quasi-experimental design and cluster sampling procedures. Therefore, the generalisability of the findings beyond this institutional context should be interpreted cautiously. The qualitative component also involved only six participants from the experimental group, thereby limiting representation of broader student experiences, including those of learners in the control group. Future studies should involve more diverse participants across varying socioeconomic, cultural, and achievement backgrounds in order to examine more comprehensively how RME influences CTNC development under different educational conditions.

Methodologically, although the mixed-methods design enriched interpretative depth through the integration of quantitative and qualitative evidence, the quasi-experimental framework remains subject to limitations in internal validity, particularly regarding the control of external variables. Future research should therefore employ more rigorous experimental approaches, such as randomised controlled trials or Solomon four-group designs, to strengthen causal inference. Longitudinal investigations across multiple mathematical domains are also necessary to examine the sustainability and transferability of CTNC development over time. In addition, further qualitative inquiry into teachers' pedagogical roles, classroom interaction patterns, and implementation challenges would provide a more comprehensive understanding of how RME functions within diverse instructional contexts.

Finally, the findings indicate that contextual factors including institutional support, curriculum structures, classroom culture, and students' readiness for collaborative learning substantially influence the effectiveness of RME implementation. Although students in the present study generally responded positively to RME instruction, outcomes may differ across educational settings characterised by distinct structural and sociocultural conditions. Accordingly, future efforts to develop and scale RME-based programmes should consider the broader readiness of schools and educational systems in order to support sustainable, context-sensitive, and equitable implementation. Overall, despite its limitations, this study provides robust preliminary evidence that RME constitutes a promising pedagogical pathway for fostering CTNC and supporting the development of critical numeracy competencies essential for participation in twenty-first-century society.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been covered completely by the authors.

References

Aizikovitsh-Udi, E., & Cheng, D. (2015). Developing Critical Thinking Skills from Dispositions to Abilities: Mathematics Education from Early Childhood to High School. *Creative Education*, *06*(04), 455–462. https://doi.org/10.4236/ce.2015.64045

Ary, D., Jacobs, L.C., Sorensen, C., Razavieh, A. (2019). Introduction to Research in Education (8th ed.). *Cengage Learning*, *11*(1), 1–14.

Butler, H. A., Pentoney, C., & Bong, M. P. (2017). Predicting real-world outcomes: Critical thinking ability is a better predictor of life decisions than intelligence. *Thinking Skills and Creativity*, *25*, 38–46. https://doi.org/10.1016/j.tsc.2017.06.005

Charmaz, K. (2014). Constructing Grounded Theory - Kathy Charmaz - Google Books. *Google Scholar*.

Creswell John and Creswell David. (2023). Research Design, Qualitative, Quantitative and Mixed Methods Approaches. SAGE Publications,Inc. (Vol. Sixth Edit, pp. 1–382).

Creswell, J.W , Poth, C. N. (2018). Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Public Administration.

Facione, P. A. (2015). Critical Thinking: What It Is and Why It Counts. Insight Assessment. *Insight Assessment*, *5*(1), 1–23.

Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics Fifth Edition*. *SAGE edge* (p. 1561).

Geiger, V., Forgasz, H., & Goos, M. (2015). A critical orientation to numeracy across the curriculum. *ZDM - International Journal on Mathematics Education*, *47*(4), 611–624. https://doi.org/10.1007/s11858-014-0648-1

Gelerstein, D., Río, R. del, Nussbaum, M., Chiuminatto, P., & López, X. (2016). Designing and implementing a test for measuring critical thinking in primary school. *Thinking Skills and Creativity*, *20*, 40–49. https://doi.org/10.1016/j.tsc.2016.02.002

Gravemeijer, K. (1994). *Developing realistic mathematics education*. *Faculty of Sciences, Freudenthal Institute*. Retrieved from http://www.cdbeta.uu.nl/tdb/fulltext/199503-terwel2.pdf http://www.fisme.science.uu.nl/toepassingen/20014/

Hammond, J., & Gibbons, P. (2005). Putting scaffolding to work: The contribution of scaffolding in articulating ESL education. *Prospect: An Australian Journal of …*, *20*(1), 6–30. Retrieved from http://www.ameprc.mq.edu.au/resources/prospect/V20_N1_2005%5Cnhttp://dialnet.unirioja.es/servlet/articulo?codigo=2437173

Heuvel-Panhuizen, M. V. D. (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. Retrieved from http://link.springer.com/article/10.1023/B:EDUC.0000005212.03219.dc

Jablonka, E. (2020). Critical thinking in mathematics education. In *Encyclopedia of mathematics education* (pp. 159-163). Cham: Springer International Publishing.

Jain, P., & Rogers, M. (2019). Numeracy as critical thinking. *Adults Learning Mathematics International Journal*, *14*(1), 23–33.

Luthfiyah, I., Simamora, R. E., & Rahayu, S. W. (2024). Eksplorasi proses pemecahan masalah matematika siswa kelas 8 sekolah menengah pertama dalam menyelesaikan soal cerita. *Mathematics Education And Application Journal (META)*, *6*(1), 30–47. https://doi.org/10.35334/meta.v6i1.5427

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods*, *16*(1). https://doi.org/10.1177/1609406917733847

Pallant, J. (2020). *SPPS: survival manual: a step by step guide to data analysis using ibm spss: 7th edition* (pp. 1–361). Taylor and Francis. https://doi.org/10.4324/9781003117452

Plomp, T., & Nieveen, N. (2013). Introduction to Educational Design Research: An Introduction. *Educational Design Research*, 11–50.

Ridder, H. G., Miles, M. B., Michael Huberman, A., & Saldaña, J. (2014). Qualitative data analysis. A methods sourcebook. *Zeitschrift Fur Personalforschung*, *28*(4), 485–487.

Rizmawati, L., Simamora, R. E., & Anwar, A. (2025). Literasi dan disposisi matematis siswa kelas X sekolah menengah atas. *Mathematics Education And Application Journal (META)*, *6*(2), 103–126. https://doi.org/10.35334/meta.v6i2.6428

Schoenfeld, A. H. (2020). Reframing teacher knowledge: a research and development agenda. *ZDM - Mathematics Education*, *52*(2), 359–376. https://doi.org/10.1007/s11858-019-01057-5

Simamora, R. E., & Ramadhanta, S. A. (2024). Investigating the effects of Realistic Mathematics Education on mathematical creativity through a mixed-methods approach. *Indonesian Journal of Science and Mathematics Education*, *7*(2), 337. https://doi.org/10.24042/ijsme.v7i2.21221

Suryawan, I. P. P., Sudiarta, I. G. P., & Suharta, I. G. P. (2023). Students' critical thinking skills in solving mathematical problems: Systematic literature review. *Indonesian Journal Of Educational Research and Review*, *6*(1), 120-133.

Van den Heuvel-Panhuizen, M. (2020). Didactical Phenomenology (Freudenthal). In *Encyclopedia of mathematics education* (pp. 218-220). Cham: Springer International Publishing.

Van den Heuvel-Panhuizen, M., & Drijvers, P. (2020). Realistic mathematics education. In *Encyclopedia of mathematics education* (pp. 713-717). Cham: Springer International Publishing.

Watson, J. M., & Callingham, R. A. (2004). Statistical literacy: From idiosyncratic to critical thinking. *Curricular Development in Statistics Education. International Association for Statistical Education (IASE) Roundtable, Lund, Sweden*, 116-162.

Wijaya, T. T., Hidayat, W., Hermita, N., Alim, J. A., & Talib, C. A. (2024). Exploring contributing factors to pisa 2022 mathematics achievement: insights from indonesian teachers. *Infinity Journal*, *13*(1), 139–156. https://doi.org/10.22460/infinity.v13i1.p139-156