PROFESI PENDIDIKAN DASAR
Volume 12, Issue 2, August 2025, pp. 98-112
p-ISSN: 2406-8012, e-ISSN: 2503-3530
https://journals2.ums.ac.id/index.php/ppd

The Contribution of Ethno-Realistic Mathematics Education (E-RME) Approach to Enhancing Elementary Students' Critical Thinking Skills

Linda Indiyarti Putri[1*], Roudlotul Afifah[1], Kharisatul Istijabah[1], Arifah Nurul Ain[1], Ulia Nasroatul Umami[1], Fadila Fatmawati[1], Haniyah Nur Sakhinah[1], Lisa Nurul Qoniah[1], Roshifah Jauhari[1], Guldana A. Begimbetova[2]

[1]Universitas Wahid Hasyim, Indonesia
[2]Abai Kazakh National Pedagogical University, Kazakhstan
*Corresponding Author's email: *lindaputridunwahas.ac.id*

How to cite: Putri, L. I., Afifah, R., Istijabah, K., Ain, A. N., Umami, U. N., Fatmawati, F., Sakhinah, H. N., Qoniah, L. N., Jauhari, R., & Guldana A. Begimbetova. (2025). The contribution of ethno-realistic mathematics education (E-RME) approach to enhancing elementary students' critical thinking skills. *Profesi Pendidikan Dasar*, 98–112. https://doi.org/10.23917/ppd.v12i2.10034

Keywords:
E-RME approach; the effectiveness of mathematics instruction; elementary school students; ANCOVA

Article History:
Submitted: 2025-05-11
Revised: 2025-08-01
Accepted: 2025-08-21

Abstract

Meaningful mathematics instruction enables students to autonomously construct mathematical concepts and discern their relevance to real-world contexts. The Ethno-Realistic Mathematics Education (E-RME) approach advances this objective by embedding cultural and contextual dimensions within the curriculum, thereby fostering critical thinking and deepening conceptual mastery. This study investigated the efficacy of the E-RME approach in enhancing the critical thinking capacities of primary school pupils. A quasi-experimental design was employed, involving 308 pupils from diverse primary schools, allocated to experimental and control groups. Assessment was conducted using the Generalised Partial Credit Model (GPCM), which facilitates polytomous scoring of two-tier test items to capture nuanced responses. Data were analysed using JASP software to evaluate pupils' conceptual understanding and reasoning proficiency based on their tiered responses. The findings revealed that the E-RME approach significantly outperformed the conventional instructional model in promoting conceptual understanding and critical thinking, even after adjusting for baseline performance via ANCOVA. These results underscore that superior learning outcomes derive not only from the mathematical content delivered but also from pedagogical strategies that meaningfully integrate mathematics with pupils' cultural identities and lived experiences.

INTRODUCTION

Background of the Study

The primary aim of mathematics education at the primary level is to endow pupils with mathematically grounded life skills that are applicable across diverse real-world contexts (Helzi et al., 2024; Rabia et al., 2024). This aim is most effectively realised through pedagogy that embeds culturally and contextually relevant frameworks, enabling pupils to forge meaningful connections between abstract mathematical concepts and their lived experiences. Rabia et al. (2024) and Setio (2023) argue that such contextualised approaches not only deepen conceptual understanding but also stimulate active engagement in critical and creative thinking, thereby fostering the development of proficient problem-solvers. When instruction is anchored in pupils' local contexts and cultural heritage, it

enhances knowledge construction through purposeful mathematical visualisation, communication, and reasoning (Nagelkop et al., 2021; Zulu & Mudaly, 2023). Consequently, mathematics education must adopt contextualised pedagogical strategies to amplify the relevance, engagement, and practical utility of mathematics in pupils' daily lives (Moore, 2023; Nagelkop et al., 2021).

Despite the focus on fostering conceptual and contextual understanding in mathematics education, pupil learning outcomes continue to fall short of expectations. Research consistently demonstrates that pupils' performance in mathematics is often inferior to that in other subjects, revealing a deficient grasp of foundational mathematical concepts (Dzulfikar & Vitantri, 2017). A principal cause of this shortfall is the dominance of procedural instruction, which fails to cultivate meaningful learning experiences. Teachers frequently rely on verbal exposition, delivering content without aligning it to pupils' learning needs, leading to rote memorisation of formulas devoid of conceptual insight (Safitri 2023). Consequently, when encountering unfamiliar material or contextual problems, pupils experience confusion, misapply strategies, and develop misconceptions, with enduring adverse effects on their academic attainment (Fujimoto et al., 2025; Ramadany, 2020).

A cornerstone of effective knowledge transfer in mathematics education is the teacher's ability to deploy a range of pedagogical strategies tailored to pupils' learning needs (Priantari et al., 2020; Susanta et al., 2023). Research consistently demonstrates that interactive and contextual teaching approaches significantly outperform traditional lecture-based methods in fostering pupils' conceptual understanding (del Valle-Ramón et al., 2020; Jdaitawi et al., 2023; Palinussa et al., 2021). Nevertheless, limited conceptual understanding of mathematics persists as a formidable challenge for both pupils and educators (Rahmadani et al., 2023). The Ethno-Realistic Mathematics Education (E-RME) approach offers an innovative pedagogical framework that integrates cultural contexts, everyday experiences, and local wisdom to cultivate pupils' conceptual mastery and critical thinking skills (Prahmana, 2022). This is substantiated by prior studies, which confirm that the E-RME approach markedly enhances pupils' mathematical critical thinking abilities, thereby elevating learning outcomes (Lubis et al., 2023; Yuliani et al., 2023). Consequently, the E-RME approach provides a robust pedagogical alternative that not only fosters mathematical proficiency but also challenges entrenched stereotypes surrounding mathematics (Boadu & Bonyah, 2024).

The Ethno-Realistic Mathematics Education (E-RME) approach substantially advances the realisation of several Sustainable Development Goals (SDGs), particularly SDG 4 (Quality Education), SDG 10 (Reduced Inequalities), and SDG 17 (Partnerships for the Goals) (Sass et al., 2023). By championing inclusive, culturally relevant, and high-quality education, and fostering collaboration and the dissemination of best practices, E-RME makes a pivotal contribution to global efforts towards sustainable development and an equitable future. This resonates with the arguments of Georgeson et al. (2017) and Johnson et al. (2022), who assert that pedagogical approaches embedding social and cultural contexts are indispensable for mitigating inequalities and achieving the inclusive, equitable, and high-quality education envisioned by the SDGs. Similarly, Agusdianita et al. (2024) demonstrate that integrating Problem-Based Learning (PBL) with a differentiated approach, especially when enriched with local cultural elements, markedly enhances pupils' learning outcomes and aligns with global educational objectives, including the SDGs.

Problem of The Study

The Ethno-Realistic Mathematics Education (E-RME) approach is widely regarded for its effectiveness in diverse school-based learning environments, yet rigorous empirical research evaluating its impact at the primary school level using advanced quantitative methods remains limited. In particular, studies employing advanced statistical tools, such as the Generalised Partial Credit Model (GPCM) and Analysis of Covariance (ANCOVA), are rare. This gap in evidence hinders a full understanding of how the E-RME approach not only boosts pupils' mathematical knowledge but also strengthens hydroxide their reasoning abilities and capacity to provide logical explanations when addressing contextual, problem-based tasks.

To bridge this research gap, this study employs a two-tier diagnostic test instrument, well-established for its ability to assess pupils' higher-order thinking skills, including their ability to offer conceptually sound solutions and justify their responses (Chandrasegaran et al., 2007; Rintayati et al.,

2020; Yusrizal & Halim, 2017). However, previous studies have not yet fully integrated this tool within the E-RME framework alongside psychometric analysis to evaluate critical thinking in culturally relevant contexts. Therefore, a focused investigation is necessary to clarify how the E-RME approach supports the development of pupils' critical thinking skills and to develop reliable methods for accurately measuring and analysing these outcomes at the primary level.

Research's State of the Art

Mathematics education has experienced a significant paradigm shift in recent years, moving away from a purely procedural focus towards a more contextual, meaningful, and conceptually grounded approach. Research has shown that the contextual learning model, which highlights the relationship between abstract mathematical ideas and real-world experiences, enhances student engagement and fosters critical thinking (Nagelkop et al., 2021; Setio, 2023). Context-based learning also supports students in constructing knowledge through activities such as communication, mathematical reasoning, and visualisation—activities that are directly relevant to their everyday lives (Moore, 2023; Zulu & Mudaly, 2023).

In this context, the Ethno-Realistic Mathematics Education (E-RME) approach emerges as an extension of Realistic Mathematics Education (RME), incorporating elements of ethnomathematics—specifically, the integration of local cultural values into mathematics instruction. By using cultural contexts as cognitive bridges, E-RME not only strengthens students' conceptual understanding but also develops their critical thinking skills (Lubis et al., 2023; Prahmana, 2022). These culturally embedded contexts serve as meaningful entry points, enabling students to connect abstract mathematical concepts with their lived experiences. Furthermore, the E-RME approach aligns with constructivist principles, which view students as active participants in the process of knowledge construction through social interaction and experiential learning.

A growing body of research underscores the effectiveness of the Ethno-Realistic Mathematics Education (E-RME) approach in enhancing students' attitudes toward mathematics, developing their critical thinking skills, and deepening their conceptual understanding (Westley, 2024; Yuliani et al., 2023). E-RME holds significant potential to contribute to the Sustainable Development Goals (SDGs), particularly SDG 4 on quality education, by offering instruction that is contextualised, inclusive, and culturally responsive. This approach fosters active student engagement while accounting for their sociocultural contexts, thereby promoting equitable and socially just educational outcomes (Johnson et al., 2022). Similarly, Georgeson et al. (2017) argue that embedding social and cultural dimensions within instructional practices can effectively address educational disparities. Consequently, E-RME may be viewed as a practical manifestation of the principle of equity in education, as it provides learning experiences that are relevant and meaningful, reflecting students' lived realities. While explicit research linking E-RME directly to the SDGs remains limited, its core characteristics align closely with the principles and objectives of sustainable education, as articulated in global educational discourse.

Gap Study & Objective

Previous research has demonstrated that E-RME enhances students' critical thinking skills and deepens their conceptual understanding of mathematics (Hariastuti et al., 2022; Prahmana, 2022; Priyambodo et al., 2023). However, significant gaps remain in the literature, particularly in relation to the application of rigorous experimental methodologies to empirically validate the effectiveness of E-RME at the elementary school level. Most existing studies have not fully assessed the efficacy of E-RME using cognitive assessment tools such as two-tier tests, which are specifically designed to investigate students' logical reasoning and argumentation within realistic, culturally embedded learning contexts.

Furthermore, there is a notable deficiency in studies that employ advanced psychometric analysis techniques, such as the Generalised Partial Credit Model (GPCM), to assess students' depth of comprehension and their ability to address contextual problems rooted in E-RME. This methodological limitation restricts our capacity to objectively compare culturally based mathematics instruction with conventional teaching approaches and to evaluate it in a valid and reliable manner.

In light of these limitations, this study aimed to evaluate the efficacy of the Ethno-Realistic Mathematics Education (E-RME) approach in enhancing critical thinking skills among elementary

school students. Specifically, it utilised a two-tier testing framework alongside statistical analyses, including GPCM and ANCOVA, to compare the E-RME approach with conventional lecture-based instruction. By addressing both theoretical and methodological shortcomings, this research contributes to the development of culturally responsive mathematics education that prioritises the cultivation of higher-order thinking skills from an early age.

METHOD

Type and Design

This study was conducted in five elementary schools in Semarang City during the second semester of the academic year. It employed a quantitative methodology within a quasi-experimental design embedded in a comparative framework. The practical constraints of achieving complete randomisation in an elementary school setting necessitated the use of a quasi-experimental design, which enabled a systematic evaluation of the effectiveness of the instructional interventions. Purposive sampling was used to select research participants based on the extent of alignment between school characteristics and teachers' preparedness to implement the E-RME-based instructional approach.

The comparative experimental design involved two groups: a control group receiving instruction through a traditional lecture method, and an experimental group taught using the E-RME approach. A purposefully developed two-tier assessment tool was used to evaluate fourth-grade students' critical thinking skills in the measurement topic. Pretests and posttests were administered to both groups to assess the impact of the instructional intervention. The two-tier instrument evaluated not only the accuracy of students' responses but also their underlying logical reasoning, offering a more comprehensive insight into their conceptual understanding. This design facilitated a systematic comparison of learning outcomes associated with the two teaching approaches.

This study employed the Generalised Partial Credit Model (GPCM) to evaluate students' responses to the test instrument, and Analysis of Covariance (ANCOVA) to compare posttest outcomes between groups while controlling for pretest scores. The GPCM, which analysed students' responses across the two tiers of the assessment, enabled a nuanced examination of their conceptual understanding. In contrast, ANCOVA facilitated the isolation of the instructional intervention's effect by adjusting for initial differences in student ability. This integrated analytical approach comprehensively assessed the efficacy of E-RME instruction in comparison to traditional teaching methods, thereby enhancing the internal validity and generalisability of the research findings.

Data and Data Sources

The research was conducted across five elementary schools in Semarang City. The study adopted a Nonequivalent Control Group Design, a quasi-experimental framework in which both the experimental and control groups underwent pretesting and posttesting (Jdaitawi et al., 2023; Sala & Gobet, 2017). Prior to data collection, a preliminary phase was conducted, involving structured interviews with teachers to assess their understanding and implementation of the E-RME-based instructional approach, particularly in relation to classroom practices and assessment instruments utilised. Sampling was conducted using a purposive sampling strategy. The final research sample comprised 308 fourth-grade students, with instruction focused on the measurement topic. Participants were divided into two structured groups: an experimental group (n = 152) and a control group (n = 156). The framework of the research design is summarised in Table 1.

Table 1. Structure of the Non-Equivalent Control Group Research Design

Group	Pretest	Treatment	Postest
Experiment	O_1	E-RME	O_2
Control	O_1	Traditional	O_2

Source : (Hastjarjo, 2019)

Data Collection Technique

Data collection was conducted through a cognitive test administered via a purpose-developed two-tier assessment instrument, specifically designed to evaluate fourth-grade elementary students' critical thinking skills within the domain of measurement. The instrument comprised a two-tier test structured to measure both conceptual understanding and reasoning processes. Tier 1 required students to provide a final answer, assessing their grasp of measurement concepts, while Tier 2 evaluated the logical reasoning underpinning their responses, capturing critical thinking dimensions such as interpretation, analysis, evaluation, and reasoning. To ensure content validity, a detailed blueprint was created to align each test item with these four critical thinking indicators.

Prior to the main administration, the instrument underwent rigorous pilot testing to establish validity and reliability. Item validity was determined using item-total correlations, while reliability was assessed through Cronbach's Alpha and McDonald's Omega coefficients. Of the 12 items initially developed, 10 demonstrated acceptable validity (item-total correlation coefficients exceeding 0.30) and were retained for the primary data collection phase. Two items (A1 and A12) were excluded due to low or negative correlation values. Reliability analysis yielded internal consistency coefficients of 0.68 (Cronbach's Alpha) and 0.69 (McDonald's Omega), indicating adequate reliability for group-level comparisons. Only items meeting these established thresholds were included in the final test.

The test was administered in two phases: a pretest before the instructional intervention and a posttest following the completion of instructional sessions in both the experimental and control groups. Testing procedures were standardised across all participating schools, with assessments conducted under strict supervision to maintain data integrity and authenticity.

Data Analysis

The data analysis process commenced with a systematic data cleaning procedure to ensure the completeness and consistency of pretest and posttest responses. Prior to primary analysis, all datasets underwent rigorous screening to identify and remove invalid or incomplete responses, including instances of multiple selections (where only one was permitted) or unanswered items. The cleaned data were subsequently transformed and modelled using statistical techniques aligned with the study's objectives. Descriptive statistics were employed to examine score distributions within each group, incorporating measures such as mean, standard deviation, and coefficient of variation. To deepen insights into students' critical thinking skills, a polytomous scoring framework based on the Generalised Partial Credit Model (GPCM) was utilised. This model proved particularly effective for the two-tier test format, as it enabled differentiation of conceptual understanding levels beyond binary correct/incorrect categorisations. By integrating both selected answers and accompanying justifications, the GPCM provided a nuanced assessment of students' reasoning processes (Chandrasegaran et al., 2007).

Inferential analysis was conducted to evaluate the relative efficacy of the E-RME approach compared to conventional instruction. Posttest performance differences between groups were analysed using either independent samples t-tests or Mann–Whitney U tests, depending on whether the data met normality assumptions. Within-group changes (pretest to posttest) were assessed via paired t-tests. To strengthen validity, Analysis of Covariance (ANCOVA) was applied to statistically control for pretest disparities, thereby enhancing the precision of intervention effect interpretation. All analyses were performed using the latest version of JASP statistical software, ensuring methodological rigour and reproducibility.

Given the study's comparative quasi-experimental design, two primary hypotheses were tested. First, it was hypothesised that both groups would demonstrate statistically significant improvements in scores from pretest to posttest, reflecting the developmental impact of the instructional approach over time. Second, it was hypothesised that the experimental group (E-RME) would achieve significantly higher posttest scores than the control group (conventional instruction), even after accounting for baseline ability levels. These hypotheses were formulated to determine whether the E-RME instructional model could produce a measurable and educationally meaningful enhancement in students' critical thinking skills relative to traditional teaching methods.

RESULTS

Developed to evaluate critical thinking within the domain of measurement, the two-tier cognitive assessment instrument aligns with the fourth-grade mathematics curriculum's learning objectives and the foundational principles of the Ethno-Realistic Mathematics Education (E-RME) framework. The first tier measures students' conceptual understanding through their final answers, while the second tier systematically examines their critical thinking via logical reasoning processes. This study operationalised critical thinking through four indicators—reasoning, interpretation, analysis, and evaluation—adapted from well-established frameworks in mathematics education. Table 1 details the scoring criteria for the two-tier items, designed to evaluate both answer accuracy and the depth of reasoning demonstrated. Table 2 further elaborates on the rubric for the item-based assessment, which holistically appraises not only the correctness of responses but also the conceptual coherence underpinning students' reasoning.

A tiered scoring system was applied: full marks (4/4) were awarded for wholly correct answers accompanied by robust logical justification, with scores decreasing incrementally based on the presence of errors or gaps in either component. This dual-focus methodology facilitates a nuanced evaluation of learners' conceptual mastery and analytical rigour, ensuring a comprehensive understanding of their cognitive engagement with the assessment tasks. This study evaluated students' learning outcomes with a focus on higher-order thinking skills (HOTS), specifically critical thinking within the cognitive domain. The two-tier assessment instrument was purposefully developed to measure students' critical thinking by evaluating both the accuracy of their final answers (Tier 1) and the quality of the reasoning underpinning those responses (Tier 2). A criterion-referenced scoring system, employing a 1–4 scale (see Table 2), enabled a granular assessment of conceptual understanding and logical reasoning. This approach aligns with established frameworks for evaluating HOTS, which require students to not only provide precise answers but also justify their reasoning through the application of relevant concepts and contextual knowledge. Table 2 outlines the evaluation criteria used to assess students' numeracy skills.

Table 2. Scoring guidelines

Tier 1	Tier 2	Score
Correct	Correct	4
Correct	False	3
False	Correct	2
False	False	1

As part of the process of evaluating item quality, a pilot study of limited scale was conducted with a sample of 228 fifth-grade students. Participants were selected on the basis of their cognitive and conceptual preparedness to engage meaningfully with the test items. This pilot phase yielded sufficiently representative data to support preliminary assessments of item validity and internal consistency reliability. Statistical analyses were performed using Pearson's correlation coefficient to evaluate item validity and Cronbach's alpha (α) to estimate reliability. All procedures were executed via JASP software, ensuring methodological rigour and precision in the analysis (See Table 3).

Table 3. Results of Item Validity and Reliability Analysis

	Item-rest correlation	**if item dropped**	
		Cronbach's α	**McDonald's ω**
A2	0.4348	0.611	0.625
A3	0.1052	0.668	0.676
A4	0.4442	0.606	0.616
A5	0.4325	0.610	0.621
A6	0.2752	0.639	0.653
A7	0.4361	0.610	0.623
A8	0.4800	0.602	0.613

A9	0.4453	0.612	0.621
A10	0.2971	0.635	0.648
A11	0.2380	0.646	0.657
A1	-0.0173	0.680	0.691
A12	-0.0316	0.682	0.693

The validity of the instrument was evaluated using an empirical validity approach based on item-total correlation analysis across the 12 test items. The results indicated that the majority of items demonstrated statistically significant positive correlations, meeting the criterion for empirical validity (correlation coefficients > 0.30). Notably, items A2, A4, A5, A7, and A9 exhibited moderately strong correlations with the total score ($r > 0.43$), underscoring their relevance and effectiveness in measuring the intended construct. However, four items—A1, A10, A11, and A12—yielded correlation coefficients below 0.30, with A1 and A12 displaying negative values. These findings suggest that these items had weak or negligible associations with the overarching construct being assessed. To enhance the instrument's psychometric quality, items A1 and A12 were excluded from subsequent analyses.

Reliability was assessed using Cronbach's Alpha and McDonald's Omega coefficients. Cronbach's Alpha values ranged from 0.60 to 0.68, while McDonald's Omega estimates showed marginally higher internal consistency, ranging from 0.61 to 0.69. Although these metrics fell within acceptable thresholds for group-level comparisons, further refinement through the removal of weaker items was deemed necessary to improve reliability. Following the exclusion of items A1 and A12, the remaining items were judged both valid and reliable for inclusion in the main data collection phase.

The validated items were subsequently utilised to evaluate students' learning outcomes via pretest and posttest administrations. Descriptive statistical analysis revealed a marked improvement in performance among students in the E-RME group compared to those receiving conventional lecture-based instruction. Specifically, the E-RME group demonstrated a significantly higher mean posttest score (M = 24.238, SD = 4.161) than the control group (M = 17.046, SD = 5.521), indicating the pedagogical efficacy of the E-RME approach in fostering critical thinking and conceptual understanding (See Table 4).

Table 4. Descriptives Statistics

	N	Mean	SD	SE	Coefficient of variation
Experiment pre test	151	17.026	6.802	0.554	0.399
Experiment post test	151	24.238	4.161	0.339	0.172
Control pre test	153	15.902	7.046	0.570	0.443
Control post test	153	17.046	5.521	0.446	0.324

Of the 308 students initially enrolled in the study—151 in the experimental group and 153 in the control group—four had incomplete pretest or posttest data. Consequently, data from 304 students were included in the final analysis. Prior to conducting parametric statistical analyses, pretest and posttest scores were subjected to assumption testing to confirm their suitability. This involved evaluating the homogeneity of variance and normality for each group. This section presents the outcomes of the data analysis, including descriptive statistics, tests for homogeneity and normality, paired-samples t-tests for within-group comparisons, independent t-tests for between-group comparisons of posttest scores, and an Analysis of Covariance (ANCOVA) to adjust for pretest scores as a covariate. These analyses aimed to assess the efficacy of the Ethno-Realistic Mathematics Education (E-RME) instructional approach, which served as the primary intervention in the study.

The experimental group's posttest scores demonstrated a normal distribution ($W = 0.983$, $p = 0.053 > 0.05$), as confirmed by the Shapiro-Wilk test for paired samples. In contrast, the control group's posttest scores exhibited a significantly non-normal distribution ($W = 0.849$, $p < 0.001$). This deviation was attributed to the wide range of student performance, with some students achieving notably low scores and others exceptionally high ones (Azzalini & Capitanio, 2018). Such patterns are commonly

observed in groups receiving conventional instruction without specialised pedagogical support (Mukherjee & Bhonge, 2025). The absence of a systematic, conceptually grounded teaching approach, such as E-RME, which fosters contextual understanding and progressive knowledge development, likely contributed to these irregularities. The results of the normality tests for both groups are detailed in Table 5.

Tabel 5. Normality Test Results (Shapiro–Wilk)

Experiment pre test	-	Experiment post test	0.983	0.053
Control pre test	-	Control post test	0.849	< .001

Note. Significant results indicate a deviation from normality.

To evaluate the impact of the intervention, paired-sample comparisons were undertaken to examine the differences between pretest and posttest scores within each group, determining the extent of any significant improvements. In the experimental group, a statistically significant increase in scores was observed from pretest to posttest ($t = -14.048$, $p < 0.001$), with a substantial effect size (Cohen's $d = -1.143$). This indicates a pronounced positive impact of the intervention on learning outcomes, consistent with findings from prior research (Agusdianita et al., 2024; Helzi et al., 2024). By contrast, the control group also exhibited a statistically significant difference between pretest and posttest scores (Wilcoxon = 3353.000, $p = 0.001$), but the effect size was notably smaller (Cohen's $d = -0.186$). This suggests that, while some progress was made, the improvement was considerably less substantial than that observed in the experimental group. The complete results of this analysis are detailed in Table 6.

Table 6. Results of the Paired Samples T-Test Analysis

Measure 1	Measure 2	Test	Statistic	z	df	P	Effect Size	SE Effect Size
Experiment pre test	- Experiment post test	Student	-14.048		150	< .001	-1.143	0.113
		Wilcoxon	653.000	-9.303		< .001	-0.882	0.094
Control pre test	- Control post test	Student	-2.300		152	0.023	-0.186	0.078
		Wilcoxon	3353.000	-3.179		0.001	-0.311	0.097

Note. For the Student's t-test, the effect size is reported using Cohen's *d* . For the Wilcoxon test, the effect size is expressed as the matched rank biserial correlation.

An independent samples t-test was performed on the pretest scores to evaluate the equivalence of prior knowledge between the experimental and control groups at the study's outset. The analysis revealed a small effect size (Cohen's $d = -0.171$) and no statistically significant difference between the groups ($t = -1.499$, $df = 305$, $p = 0.135$). This suggests that the baseline mathematical competencies of both groups were broadly comparable before the intervention.

In contrast, analysis of the post-test scores identified a statistically significant difference between the groups ($t = -12.815$, $df = 302$, $p < 0.001$), accompanied by a large effect size (Cohen's $d = -1.470$). These findings indicate that the Ethno-Realistic Mathematics Education (E-RME) instructional approach substantially improved students' learning outcomes. The Brown–Forsythe test revealed a violation of the homogeneity of variances assumption for the post-test data ($p < 0.05$). Consequently, applying Welch's correction may be appropriate to ensure the robustness of the between-group comparison. A summary of the findings from the between-group comparisons is presented in Table 7.

Table 7. Results of Pre-test and Post-test Comparisons Between Groups

Independent Samples T-Test					
	t	df	p	Cohen's d	SE Cohen's d
Pretest	-1.499	305	0.135	-0.171	0.115
Posttest	-12.815	302	< .001	[a] -1.470	0.142

Note. Independent samples t -test results are reported.

[a] Brown-Forsythe test is significant (p < .05), indicating a violation of the assumption of equal variances

Given that the posttest data deviated from a normal distribution, a non-parametric alternative was employed. The Mann–Whitney *U* test revealed a statistically significant difference in posttest performance between the experimental and control groups ($p < 0.001$). An effect size of –0.721 indicates a substantial treatment effect, supporting the conclusion that the instructional approach implemented in the experimental group was markedly more effective than the traditional method used in the control group.

Table 8. Results of ANCOVA Analysis

ANCOVA - post tes

Cases	Sum of Squares	df	Mean Square	F	p	η^2
Pretest	1724.308	1	1724.308	94.267	< .001	0.154
Model	3495.730	1	3495.730	191.111	< .001	0.313
Residuals	5505.789	301	18.292			

Note. Type III Sum of Squares

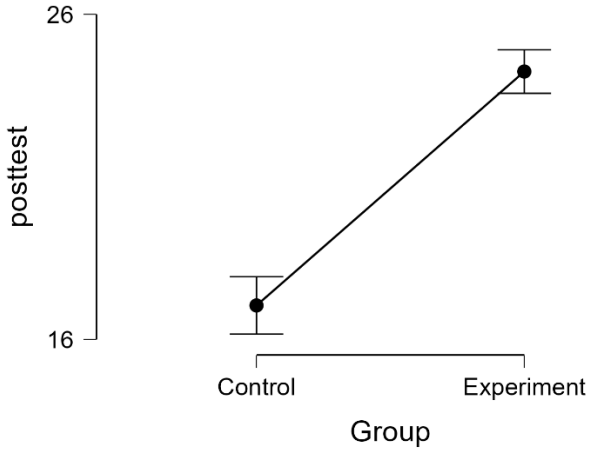
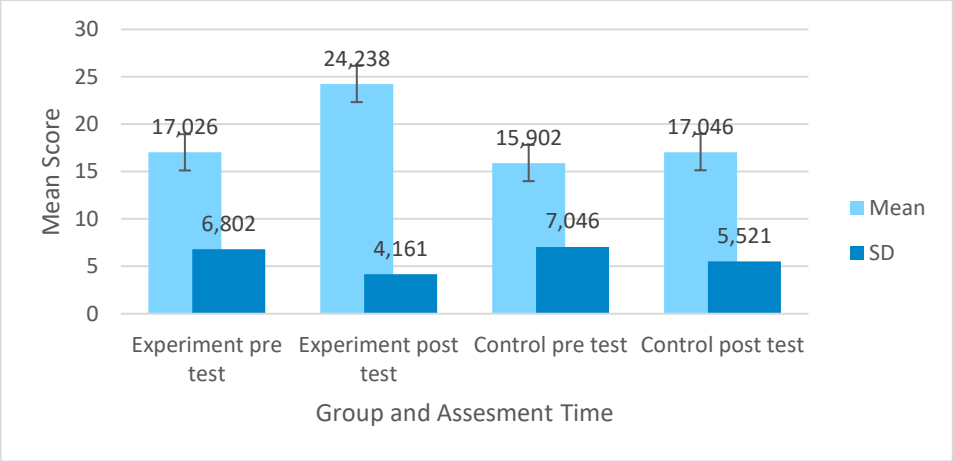


Figure 1. Mean Post-test Scores with Error Bars for Experimental and Control Groups

Based Table 8, an Analysis of Covariance (ANCOVA), incorporating pretest scores as a covariate, was employed to examine differences in posttest outcomes between the experimental and control groups, thereby substantiating prior evidence of the E-RME approach's efficacy in cultivating students' critical thinking skills. Following adjustment for baseline differences in prior knowledge, the results revealed a statistically significant disparity between the two groups ($p < 0.001$). This aligns with Wright's (2006) assertion that ANCOVA provides a more precise estimation of intervention effects in pretest–posttest designs through its capacity to account for baseline variations in student performance. Figure 1 illustrates that the experimental group consistently achieved higher mean posttest scores compared to the control group. The non-overlapping standard error bars further suggest a statistically significant difference between the groups (Novikasari, 2020). This visual representation underscores the positive impact of the Ethno-Realistic Mathematics Education (E-RME) instructional approach on student learning outcomes, while also affirming its empirical validity as an effective educational intervention.

DISCUSSIONS

This study provides robust evidence that the Ethno-Realistic Mathematics Education (E-RME) approach, which synthesises realistic mathematics instruction with local cultural contexts, demonstrates a statistically significant enhancement in cognitive learning outcomes for elementary students. Compared to the control group, which received conventional lecture-based instruction, the experimental group—exposed to E-RME-based pedagogy—exhibited superior, more consistent, and evenly distributed learning outcomes. These results are consistent with constructivist learning theory, which posits that learners construct understanding through active engagement in meaningful contexts and first-hand experiential learning (Juandi & Tamur, 2021; Susanti et al., 2023). E-RME makes



education more relevant and deeper by incorporating social and cultural realities into pedagogy, bridging the gap between students' lived realities and abstract mathematical concepts.

Figure 2. Comparison of Mean Pre-test and Post-test Scores Experimental and Control Groups

Figure 2 reveals that students taught through the Enhanced Realistic Mathematics Education (E-RME) approach achieved notably higher learning outcomes than those instructed via conventional lecture-based methods. Regarding performance consistency, the E-RME group exhibited a coefficient of variation of 0.172, considerably lower than the 0.324 observed in the lecture group. This reduced variability indicates that student performance in the E-RME group was more uniform and stable. Consequently, the E-RME approach not only demonstrated greater effectiveness in enhancing academic achievement but also provided more consistent outcomes across learners. These results align with prior studies by Limbago-Bastida and Bastida (2022) and Rahmayanti et al. (2020), which underscored the value of strategic intervention materials in boosting student performance. Such findings reinforce the argument for contextually tailored instructional strategies, like E-RME, as pivotal in fostering cognitive development.

Further validating these conclusions, an analysis of covariance (ANCOVA) established that the observed improvements in post-test scores were attributable to the E-RME intervention, rather than pre-existing differences in student ability. The robustness of ANCOVA in this context is corroborated by Valente and MacKinnon (2017) and Qi et al. (2022), who affirm its precision and minimal bias in evaluating treatment effects within pretest-posttest control group designs. In practice, students in the E-RME group benefited from a curriculum that integrated foundational knowledge with real-world applications. This experiential learning approach markedly enhanced their capacity to tackle contextual problems in assessments (Badawi et al., 2021; Jonas, 2018). These outcomes confirm that the superior performance of the E-RME group stemmed directly from the instructional strategy, rather than initial disparities in aptitude. As a constructivist and contextually grounded model, E-RME fosters meaningful learning experiences, empowering students to build knowledge effectively (Juandi & Tamur, 2021; Susanti et al., 2023; Wildaniati, 2019).

In this study, the application of Analysis of Covariance (ANCOVA) adjusts for initial pretest scores, providing a robust evaluation of the intervention's efficacy. Consequently, the significant

improvements in post-test scores observed in the experimental group can be confidently attributed to the Enhanced Realistic Mathematics Education (E-RME) approach, rather than pre-existing differences in student ability. This reinforces the conclusion that E-RME is both statistically robust and pedagogically transformative. Additionally, the use of a meticulously developed and validated two-tier instrument, analysed through the Generalized Partial Credit Model (GPCM), offers compelling evidence of E-RME's effectiveness in cultivating students' logical reasoning and critical thinking skills (Kurnia, 2019). This tiered assessment model captures the depth of students' understanding by evaluating both the accuracy of their responses and the reasoning processes underpinning them (Muraki, 1997).

These findings align with a growing body of research that underscores the value of contextual learning grounded in local wisdom for fostering higher-order thinking skills (Smith et al., 2020; Lubis et al., 2023; Prahmana, 2022). The E-RME approach leverages cultural diversity as a rich pedagogical resource, enabling students to forge meaningful connections between mathematical concepts and their cultural identities. By embedding local contexts—such as traditional measurement practices, daily routines, enduring community traditions, and embedded social values—E-RME creates authentic learning experiences. This approach resonates with Vygotsky's sociocultural theory, which asserts that learning is most effective when anchored in students' sociocultural environments (Lovorn & Summers, 2013).

In addition, the E-RME approach substantially advances culturally responsive pedagogy by not only embracing students' cultural diversity but also fostering deeper engagement and sustained commitment to the learning process. This inclusive instructional framework strengthens students' affective connections to mathematical content, thereby enhancing their intrinsic motivation to learn. In doing so, E-RME transcends conventional pedagogical models, functioning as a form of cultural affirmation within primary education that aligns closely with the principles of culturally sustaining pedagogy (Faragher et al., 2016; Gervasoni & Peter-Koop, 2020). However, this study presents several notable limitations. Foremost, the successful implementation of E-RME critically depends on teachers' capacity to design and deliver culturally relevant instructional materials. Many educators may lack adequate training and resources necessary to effectively adapt learning materials to ethnomathematical contexts. Additionally, the study's geographically constrained scope—limited to a single metropolitan area—may not fully capture the rich cultural diversity present across Indonesia. Consequently, caution is warranted when generalizing these findings beyond the studied context. Further research is essential to evaluate the efficacy of E-RME across diverse cultural settings, including rural and indigenous communities.

Moreover, this investigation did not explore important non-cognitive dimensions such as collaborative problem-solving, mathematical communication, or students' cultural mathematical literacy, all of which constitute critical 21st-century competencies. Future studies could broaden the evaluative framework by incorporating authentic, performance-based assessments that measure social skills, affective attitudes toward mathematics, and cross-contextual numeracy. Given the ongoing digital transformation of education, examining the effectiveness of E-RME within technology-enhanced, online, or hybrid learning environments also represents a timely and relevant avenue for further inquiry.

The originality of this study is underscored by its innovative integration of the E-RME approach, the application of a two-tier assessment instrument analysed through the Generalized Partial Credit Model (GPCM), and its focus on cultivating elementary students' critical mathematical thinking within the cultural context of Indonesia. This research addresses a significant methodological gap in ethnomathematics scholarship, which has historically lacked robust quantitative validation, while simultaneously providing both theoretical and empirical foundations for advancing a more inclusive, reflective, and culturally responsive mathematics curriculum. The findings highlight the imperative for comprehensive teacher preparation, the development of culturally relevant instructional materials, and the intentional incorporation of local values into national curriculum frameworks. By doing so, this study ensures that mathematics education remains deeply connected to students' sociocultural realities, thereby fostering meaningful and contextually grounded learning experiences.

CONCLUSION

The E-RME approach, by embedding local cultural contexts within its instructional design, not only fosters students' critical thinking and higher-order cognitive skills but also represents a transformative contribution to mathematics education. The integration of contextual and ethnomathematical elements, coupled with a two-tier assessment instrument analysed through the Generalised Partial Credit Model (GPCM) and supported by Analysis of Covariance (ANCOVA), offers a robust and comprehensive method for evaluating both students' final responses and their underlying reasoning processes—dimensions often neglected in conventional assessments. This methodological innovation, which bridges cognitive theory with culturally responsive pedagogy, establishes E-RME as a promising framework for nurturing critical thinking literacy in primary education. However, the study's focus on a specific geographical area and grade level limits the generalisability of its findings. Future research should explore the application of E-RME across diverse educational levels, cultural contexts, and mathematical domains to further substantiate its efficacy. These findings underscore the imperative for educators and policymakers to adopt instructional approaches that transcend procedural learning, embracing inclusivity, reflecting students' lived experiences, and fostering deep conceptual understanding and critical thinking from an early age.

REFERENCES

Agusdianita, N., Danim, S., Susanta, A., Dwi Shofiya Maghfira Izzania, R., & Irmayanti, M. (2024). Problem-based learning materials integrated with differentiated approaches to enhance elementary school students' learning outcomes. *Profesi Pendidikan Dasar*, 161-182. https://doi.org/10.23917/ppd.v11i2.6441

Azzalini, Adelchi., & Capitanio, Antonella. (2018). *The skew-normal and related families.* Cambridge University Press.

Badawi, S., Carrubbo, L., Dragoicea, M., & Walletzký, L. (2021). An analysis of the complex education service for resilience in a multi-contextual framework. In *INTED2021 Proceedings* (pp. 5159-5168). IATED. https://doi.org/10.21125/inted.2021.1060

Boadu, S. K., & Bonyah, E. (2024). The role of philosophy of mathematics education in mathematics teacher education. *Cogent Education*, 11(1). https://doi.org/10.1080/2331186X.2024.2433832

Chandrasegaran, A. L., Treagust, D. F., & Mocerino, M. (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. *Chemistry Education Research and Practice,* 8(3), 293–307. https://doi.org/10.1039/B7RP90006F

del Valle-Ramón, D., Muñoz-Repiso, A. G. V., & Gómez-Pablos, V. B. (2020). Project-based learning through the youtube platform for teaching mathematics in primary education. *Education in the Knowledge Society,* 21, 1–9. https://doi.org/10.14201/eks.20272

Dzulfikar, A., & Vitantri, C. A. (2017). Miskonsepsi matematika pada guru sekolah dasar. *Suska Journal of Mathematics Education*, 3(1), 41. https://doi.org/10.24014/sjme.v3i1.3409

Faragher, R., Clarke, B., & Hil, J. C. (2016). *Inclusive Practices in Mathematics Education*. Research in Mathematics Education in Australasia 2012-2015, December 2017, 1–378. https://doi.org/10.1007/978-981-10-1419-2

Fujimoto, Y., Azmat, F., Presbitero, A., Thomas Homer, S., Poon, W. C., & Lim, W. M. (2025). Navigating paradoxes of responsible management education (RME): insights from the Global South. *Studies in Higher Education*, 1–18. https://doi.org/10.1080/03075079.2025.2546955

Georgeson, L., Maslin, M., & Poessinouw, M. (2017). The global green economy: a review of concepts, definitions, measurement methodologies and their interactions. *Geo: Geography and Environment*, 4(1). https://doi.org/10.1002/geo2.36

Gervasoni, A., & Peter-Koop, A. (2020). Inclusive mathematics education. *In Mathematics Education Research Journal* (Vol. 32, Issue 1). https://doi.org/10.1007/s13394-020-00315-0

Hariastuti, R. M., Budiarto, M. T., & Manuharawati. (2022). Traditional houses in ethnomathematical-thematic-connected-based mathematics learning. *International Journal of Educational Methodology*, 8(3), 535–549. https://doi.org/10.12973/ijem.8.3.535

Hastjarjo, T. D. (2019). Rancangan Eksperimen-Kuasi. *Buletin Psikologi,* 27(2), 187. https://doi.org/10.22146/buletinpsikologi.38619

Helzi, H., Minsih, M., Desstya, A., & Muhammad, M. (2024). Innovative approaches to science and social studies-centered learning with a focus on student wellbeing in elementary education. *Profesi Pendidikan Dasar*, 64–79. https://doi.org/10.23917/ppd.v11i1.6173

Jdaitawi, M., Muhaidat, F., Alsharoa, A., Alshlowi, A., Torki, M., & Abdelmoneim, M. (2023). The Effectiveness of Augmented Reality in Improving Students Motivation: An Experimental Study. *Athens Journal of Education*, 10(2), 365–380. https://doi.org/10.30958/aje.10-2-10

Johnson, J. D., Smail, L., Corey, D., & Jarrah, A. M. (2022). Using bayesian networks to provide educational implications: mobile learning and ethnomathematics to improve sustainability in mathematics education. *Sustainability (Switzerland),* 14(10). https://doi.org/10.3390/su14105897

Jonas, N. (2018). Numeracy practices and numeracy skills among adults. OECD *Education Working Papers*, No. 177, OECD Publishing, Paris. http://dx.doi.org/10.1787/8f19fc9f-en

Kurnia, A. (2019). Analisis tes kemampuan berpikir kritis matematika siswa dengan menggunakan generalized partial credit model (GPCM). *In PEDIAMATIKA: Journal of Mathematical Science and Mathematics Education* (Vol. 01, Issue 02). http://syekhnurjati.ac.id/jurnal/index.php/pmat

Limbago-Bastida, R. A. C., & Bastida, G. L. (2022). Effectiveness of strategic intervention material on the learning outcomes of students. *European Journal of Social Sciences Studies*, 7(4). https://doi.org/10.46827/ejsss.v7i4.1249

Lovorn, M., & Summers, R. (2013). An international discourse on culture in the classroom: teachers from four countries share their perspectives. *International Journal of Education*, 5(2). https://doi.org/10.5296/ije.v5i2.3097

Lubis, F. F., Nurdin, E., & Fitri, I. (2023). Pembelajaran Ethno-RME Meningkatkan Keterampilan Berpikir Kritis Matematis Siswa. *Juring (Journal for Research in Mathematics Learning),* 6(3), 277. https://doi.org/10.24014/juring.v6i3.25754

Moore, M. (2023). Multicontextual teaching and learning in postsecondary classrooms. *Collected Essays on Learning and Teaching*. https://celt.uwindsor.ca/index.php/CELT/article/view/7141

Mukherjee, H., & Bhonge, P. (2025). Assessing Skew Normality in Marks Distribution, a Comparative Analysis of Shapiro Wilk Tests. *arXiv preprint arXiv:2501.14845*. http://arxiv.org/abs/2501.14845

Nagelkop, N. D., Rosselló, M., Aranguren, I., Lado, V., Ron, M., & Toglia, J. (2021). Using multicontext approach to improve instrumental activities of daily living performance after a stroke: A case report. *Occupational therapy in health care*, *35*(3), 249-267. https://doi.org/10.1080/07380577.2021.1919954

Novikasari, I. (2020). Pre-service teacher's mathematical knowledge for teaching in problem-based learning. *JRAMathEdu (Journal of Research and Advances in Mathematics Education),* 5(2), 160–174. https://doi.org/10.23917/jramathedu.v5i2.10556

Rahmayanti, S., Subagiharti, H., & Herawati, T. (2020). Effectiveness of Learning Mathematics with Contextual Approaches. *International Journal Of Multi Science*, *1*(05), 1-5. https://multisciencejournal.com/index.php/ijm/article/view/31

Palinussa, A. L., Molle, J. S., & Gaspersz, M. (2021). Realistic mathematics education: Mathematical reasoning and communication skills in rural contexts. *International Journal of Evaluation and Research in Education,* 10(2), 522–534. https://doi.org/10.11591/ijere.v10i2.20640

Prahmana, R. C. I. (2022). Ethno-Realistic Mathematics Education: The promising learning approach in the city of culture. *SN Social Sciences*, 2(12), 257. https://doi.org/10.1007/s43545-022-00571-w

Priantari, I., Prafitasari, A. N., Kusumawardhani, D. R., & Susanti, S. (2020). Improving student critical thinking trough STEAM-PjBL learning. *Bioeducation Journal*, 4(2), 95–103. https://doi.org/10.24036/bioedu.v4i2.283

Priyambodo, P., Paidi, P., Wilujeng, I., & Widowati, A. (2023). Ethno-ECLIPSE Learning Model: The Bridge between Collaboration and Critical Thinking Skills. *Journal of Education and Learning (EduLearn)*, *17*(4), 575-588. https://doi.org/10.11591/edulearn.v17i4.20876

Qi, H., Rizopoulos, D., & van Rosmalen, J. (2022). Incorporating historical control information in ANCOVA models using the meta-analytic-predictive approach. *Research Synthesis Methods*, 13(6), 681–696. https://doi.org/10.1002/jrsm.1561

Rabia, S. F., Asrul, A., Genç, N. E., & Azizah, M. A. N. (2024). Instructional material design of project-based learning to train creative thinking skills in society 5.0 era. *Profesi Pendidikan Dasar,* 96–110. https://doi.org/10.23917/ppd.v11i2.2846

Rahmadani, M. L., Zulfah, Z., & Zulhendri, Z. (2023). Pengaruh Pendekatan Ethno-RME terhadap Kemampuan Pemahaman Konsep Matematis Siswa. *Jurnal Pendidikan MIPA*, *13*(4), 1162-1170. https://doi.org/10.37630/jpm.v13i4.1350

Ramadany, L. D. (2020). Analisis miskonsepsi siswa kelas V dalam meyelesaikan masalah bangun ruang berdasarkan gender di SD IT Mutiara Insan Sorong. *Jurnal Papeda: Jurnal Publikasi Pendidikan Dasar,* 2(1), 17–26. https://doi.org/10.36232/jurnalpendidikandasar.v2i1.406

Rintayati, P., Lukitasari, H., & Syawaludin, A. (2020). Development of two-tier multiple choice test to assess indonesian elementary students' higher-order thinking skills. *International Journal of Instruction*, 14(1), 555–566. https://doi.org/10.29333/IJI.2021.14133A

Safitri, K. D., Arifiati, B. A., & Susanti, P. (2023). Analisis miskonsepsi pembelajaran matematika pada guru dan siswa sekolah Dasar. *Jurnal Ilmiah Mandalika Education (MADU)*, *1*(2), 253-257. https://doi.org/10.36312/madu.v1i2.51

Sala, G., & Gobet, F. (2017). Does chess instruction improve mathematical problem-solving ability? Two experimental studies with an active control group. Learning and Behavior, 45(4), 414–421. https://doi.org/10.3758/s13420-017-0280-3

Sass, W., De Maeyer, S., Boeve-de Pauw, J., & Van Petegem, P. (2023). Honing action competence in sustainable development: what happens in classrooms matters. *Environment, Development and Sustainability,* 25(4), 3649–3670. https://doi.org/10.1007/s10668-022-02195-9

Setio, A. (2023). *Statistical literacy : students ' performance in solving numeracy-based statistic questions*. Atlantis Press International BV. https://doi.org/10.2991/978-94-6463-220-0_6

Smith, W. M., Beattie, H. L., Ren, L., & Heaton, R. M. (2020). The evolution of a child study assignment: a potential approach to developing elementary mathematics teachers' professional noticing. *Journal of Early Childhood Teacher Education*, 43(3), 363–388. https://doi.org/10.1080/10901027.2020.1852346

Juandi, D., & Tamur, M. (2021). Review of problem-based learning trends in 2010-2020 : A meta-analysis study of the effect of problem-based learning in enhancing mathematical problem-solving skills of Indonesian students. *Journal of Physics: Conference Series*, 1772(1). https://doi.org/10.1088/1742-6596/1722/1/012103

Susanta, A., Dewi Rahimah, Irwan Koto, Edi Susanto, Muchlis, E. E., & Mela Azizah. (2023). Enhancing elementary literacy skills through a contextualized coastal course book: a developmental study in Bengkulu, Indonesia. *Profesi Pendidikan Dasar*, 1–16. https://doi.org/10.23917/ppd.v10i3.3169

Susanti, M., Suyanto, S., Jailani, J., & Retnawati, H. (2023). Problem-Based learning for improving problem-solving and critical thinking skills: A case on probability theory course. *Journal of Education and Learning (EduLearn)*, *17*(4), 507-525. https://doi.org/10.11591/edulearn.v17i4.20866

Valente, M. J., & MacKinnon, D. P. (2017). Comparing models of change to estimate the mediated effect in the pretest–posttest control group design. *Structural Equation Modeling*, 24(3), 428–450. https://doi.org/10.1080/10705511.2016.1274657

Westley, L. (2024). How might a mastery style of teaching in mathematics lead to deeper understanding and sustained levels of attainment? *Education* 3-13, 1–14. https://doi.org/10.1080/03004279.2024.2405023

Wildaniati, Y. (2019). Developing problem-based mathematics teaching materials to facilitate problem-solving ability achievement. *Elementary: Jurnal Iilmiah Pendidikan Dasar*, *5*(1), 27-38. https://e-journal.metrouniv.ac.id/elementary/article/view/1247

Yuliani, S. B., Dewi, S. K., Ain, Z. Q., & Palupi, E. L. W. (2023). Pengembangan modul berbasis etno-RME berbalut konteks wayang kulit mahabharata pada materi himpunan untuk siswa kelas 7. *MATHEdunesa*, 12(1), 108–128. https://doi.org/10.26740/mathedunesa.v12n1.p108-128

Yusrizal, Y., & Halim, A. (2017). The effect of the one-tier, two-tier, and three-tier diagnostic test toward the students'confidence and understanding toward the concepts of atomic nuclear. *Unnes Science Education Journal*, *6*(2). https://journal.unnes.ac.id/sju/usej/article/view/15856

Zulu, M. W., & Mudaly, V. (2023). Unveiling problem-solving strategies of pre-service mathematics teachers: A visual and discursive exploration. *Eurasia Journal of Mathematics, Science and Technology Education,* 19(7). https://doi.org/10.29333/ejmste/13344