



Effects of Edu-Board Game-Assisted STAD on Elementary School Students' Mathematics Problem Solving: A Quasi-Experimental Study

¹Febby Deca Lestari, ^{1*}Hafiziani Eka Putri, ¹Teten Ginanjar Rahayu, ¹Alfiana Nurussama, ²Mariano Dos Santos

¹Elementary School Teacher Education, Universitas Pendidikan Indonesia, Bandung, Indonesia

²Instituto Católico Para A Formação De Professores (ICFP), Baucau, Timor - Leste

*Corresponding author e-mail: hafizianiekaputri@upi.edu

Received: June 2026; Revised: September 2026; Published: October 2026

Abstract

Mathematical problem-solving is an essential competency in elementary education, yet students often experience difficulties when solving problems involving Least Common Multiple (LCM) and Greatest Common Factor (GCF). This study compared changes in mathematical problem-solving performance between Grade IV students receiving Edu-Board Game-Assisted Student Teams Achievement Divisions (STAD) and those receiving Problem-Based Learning (PBL) supported by ice-cream-stick manipulatives. A quasi-experimental nonequivalent pretest-posttest design involved 60 students from two intact classes, with 30 students in each instructional condition. Mathematical problem-solving was assessed using five constructed-response problems based on four problem-solving stages. Both groups showed substantial improvement from pretest to posttest. The experimental group increased from a mean of 41.33 to 73.67, while the comparison group increased from 26.27 to 60.53. The groups differed significantly at baseline, $t(43.76) = -4.77, p < .001$. After accounting for pretest performance, the adjusted posttest difference favored the experimental condition but was not statistically significant, $B = 4.04, 95\% \text{ CI } [-0.53, 8.60], F(1, 57) = 3.14, p = .082, \text{ partial eta squared} = .052$. Supplementary analysis showed higher N-Gain in the experimental group (57.12%) than in the comparison group (46.34%), $t(58) = 3.30, p = .002, \text{ Cohen's } d = 0.851$. These findings indicate that Edu-Board Game-Assisted STAD represents a viable active approach for elementary mathematical problem solving, although its adjusted advantage over active PBL was not statistically established.

Keywords: Cooperative learning; Educational board game; Mathematical problem solving; Problem-based learning; Student teams achievement divisions

How to Cite: Lestari, F. D., Putri, H. E., Rahayu, T. G., Nurussama, A., & Santos, M. D. (2026). Effects of Edu-Board Game-Assisted STAD on Elementary School Students' Mathematics Problem Solving: A Quasi-Experimental Study. *Prisma Sains: Jurnal Pengkajian Ilmu Dan Pembelajaran Matematika Dan IPA IKIP Mataram*, 14(4), 2097–2116. <https://doi.org/10.33394/j-ps.v14i4.21336>



<https://doi.org/10.33394/j-ps.v14i4.21336>

Copyright© 2026, Lestari et al.

This is an open-access article under the [CC-BY](https://creativecommons.org/licenses/by/4.0/) License



INTRODUCTION

Mathematical problem-solving is a central competency in elementary mathematics because it requires students not only to recall procedures but also to understand a problem, select an appropriate strategy, carry out the solution, and evaluate the result. This competency becomes increasingly important in Grade IV, when students begin to encounter more complex and application-oriented topics, including Least Common Multiple (LCM) and Greatest Common Factor (GCF), that require reasoning beyond routine arithmetic procedures (Pólya, 1957). Developing problem-solving competence at this stage is important because it provides a foundation for subsequent mathematical learning and enables students to connect mathematical ideas with situations encountered in everyday life (Santos-Trigo, 2024; Siswanto & Meiliasari, 2024; Vicente et al., 2022). Evidence from Indonesian elementary education

nevertheless indicates that students continue to experience difficulties with numeracy problems, particularly those involving number operations, division, and LCM in the Minimum Competency Assessment (AKM) (Nurarini & Mariana, 2024). Similar evidence indicates that elementary students' success in mathematical problem solving is influenced by the linguistic, numerical, and representational characteristics of the problems they encounter (Vessonen et al., 2024; Wisenöcker et al., 2024). These findings suggest that strengthening mathematical problem-solving should remain an important instructional priority in elementary education.

Persistent difficulties in mathematical problem-solving are evident in students' ability to interpret information, determine appropriate solution strategies, and carry out procedures accurately. Elementary students may struggle to identify relevant information in word problems, coordinate multiple cognitive demands, and determine the mathematical procedures required to reach a solution (Huang et al., 2024; Prang et al., 2024; Sagita et al., 2023; Strohmaier et al., 2022; Vessonen et al., 2024). Such difficulties may become more pronounced when students encounter problems that differ from previously practiced examples and therefore require flexible rather than reproductive reasoning (Säfsström et al., 2024; Žakelj et al., 2025). At the instructional level, mathematics learning is also frequently characterized by teacher-centered explanation, routine assignments, and repeated exercises that offer limited opportunities for students to discuss, explore, and construct solutions actively (Asmedy, 2021; Bramantha & Meliandani, 2024). Under these conditions, students may have relatively few opportunities to articulate mathematical reasoning, compare alternative strategies, or test their understanding collaboratively. Research on teaching through problem solving emphasizes the importance of providing students with opportunities to develop, explain, compare, and reflect on solution strategies rather than focusing primarily on obtaining correct answers (Roorda et al., 2024; Stewart & Ball, 2023). Instructional media may therefore be useful when they are embedded within pedagogical approaches that encourage active participation rather than used merely as supplementary classroom materials (Damayanti et al., 2025).

One pedagogical approach that may provide such opportunities is Student Teams Achievement Divisions (STAD), a form of cooperative learning that organizes students into small teams while maintaining individual accountability through structured learning activities and assessment. Cooperative structures can encourage students to exchange explanations, contribute to shared tasks, and take responsibility for both individual and group learning (Barreto et al., 2022; Nguyen & Oanh, 2025). These features are potentially relevant to mathematical problem-solving because students need opportunities to explain strategies, examine different approaches, and justify solutions. Physical educational board games may complement this cooperative structure by providing a rule-based environment in which mathematical tasks are embedded in interactive activities. Evidence from primary education indicates that board-game interventions can support academic skills and cognitive development when they are incorporated systematically into classroom learning (Vita-Barrull et al., 2024). Previous research has also shown that mathematical board games can support learning through concrete representations, repeated practice, and interaction around mathematical content (Fathurrohman et al., 2022; Lin & Cheng, 2022; Siregar, 2021). In the present study, the term Edu-Board Game refers to a complete physical, rule-based educational game containing Grade IV mathematical problem-solving tasks. This differs conceptually from gamification, which involves adding game-design elements to a non-game activity rather than using a complete game as the learning environment (Deterding et al., 2011).

Existing empirical research provides support for examining STAD and educational games in elementary mathematics, although the evidence remains fragmented across different instructional designs and outcomes. STAD has been compared with other active learning approaches, with some studies reporting favorable gains in mathematical problem-solving outcomes (Nurmeliyanti et al., 2024). Board-game-based instruction has also been associated with improved mathematical learning and problem-solving, although some studies have relied

on one-group pretest-posttest designs or have focused on general learning outcomes rather than mathematical problem-solving specifically (Azzahra et al., 2026; Safitri, 2019). Development studies have further demonstrated the feasibility and practicality of mathematics-oriented board games, but such studies do not necessarily establish whether the resulting learning outcomes differ from those produced by alternative instructional approaches (Drestanta et al., 2025). More rigorous evidence from primary education nevertheless indicates that classroom board-game interventions can generate improvements in mathematics-related academic outcomes relative to regular instruction, although the magnitude of improvement may vary across settings and learner groups (Vita-Barrull et al., 2024). At the same time, comparisons between active pedagogies such as Problem-Based Learning (PBL) and Discovery Learning indicate that mathematical problem-solving outcomes may vary even when both conditions involve student-centered learning (Kusmiyantoro et al., 2026). Together, these findings suggest that evaluating an instructional innovation against another active comparison condition may provide a more informative test than comparing it only with passive or conventional instruction.

Despite this growing body of research, several gaps remain. STAD and educational board games have commonly been examined as separate instructional components, whereas fewer studies have investigated their integration within a structured elementary mathematics lesson. In addition, board-game interventions have often been evaluated through developmental studies, single-group designs, or comparisons that do not involve another active instructional model. Evidence from mathematics education further suggests that problem-solving performance is shaped not only by the instructional medium but also by the cognitive demands of the task, the opportunities provided for mathematical reasoning, and the extent to which students are supported in moving beyond superficial solution procedures (Säfström et al., 2024; Strohmaier et al., 2022; Vicente et al., 2022). Consequently, evidence remains limited regarding whether combining STAD with a physical, rule-based educational board game is associated with different mathematical problem-solving outcomes when compared with an alternative active approach such as PBL supported by concrete manipulatives. This distinction is important because a comparison with an active condition provides a more demanding test of the instructional approach and reduces the likelihood that any observed difference can be attributed simply to students receiving more interactive instruction. The present study therefore focuses on Edu-Board Game-Assisted STAD as a combined instructional condition rather than attempting to isolate the independent contribution of either STAD or the board game.

Against this background, the present study compared Grade IV students' mathematical problem-solving performance under two active instructional conditions: Edu-Board Game-Assisted STAD and PBL assisted by ice-cream-stick manipulatives. The instructional content was limited to LCM and GCF, allowing both groups to engage with the same mathematical domain while experiencing different combinations of pedagogical structure and learning media. The study was designed as a nonequivalent-group quasi-experiment using intact classes; therefore, the comparison was intended to examine differences in learning outcomes under the two instructional conditions rather than to establish an unequivocal causal effect of either approach. Particular attention was given to students' initial mathematical problem-solving ability because baseline differences between intact classes can influence post-intervention comparisons. The contribution of the study lies in providing evidence from an active-comparator design in which a cooperative, board-game-supported condition is evaluated against PBL with concrete manipulatives in elementary mathematics.

Accordingly, the study addressed the following research questions:

- RQ1. How did students' mathematical problem-solving performance change from pretest to posttest under Edu-Board Game-Assisted STAD and ice-cream-stick-assisted Problem-Based Learning?

- RQ2. How did students' posttest mathematical problem-solving performance differ between Edu-Board Game-Assisted STAD and ice-cream-stick-assisted Problem-Based Learning when baseline ability was taken into account?

METHOD

Research Design

This study employed a quantitative approach using a quasi-experimental nonequivalent pretest-posttest group design. The design was selected because the participants were drawn from two pre-existing Grade IV classes and were not randomly assigned at the individual level. One intact class received Edu-Board Game-Assisted Student Teams Achievement Divisions (STAD), whereas the other intact class received Problem-Based Learning (PBL) assisted by ice-cream-stick manipulatives. Both instructional conditions covered the same mathematical content, namely Least Common Multiple (LCM) and Greatest Common Factor (GCF), and were implemented over three instructional sessions with equivalent time allocation and assessment procedures. Because the study involved intact classes and each instructional condition was represented by one class, the design was intended to compare mathematical problem-solving outcomes between the two instructional conditions rather than to establish an unequivocal causal effect of either approach. The structure of the research design is presented in Table 1.

Table 1. Nonequivalent Pretest-Posttest Group Design

Group	Pretest	Instructional Condition	Posttest
Experimental	O ₁	X ₁	O ₂
Comparison	O ₃	X ₂	O ₄

In Table 1, O₁ represents the pretest administered to the experimental group, while O₂ represents the corresponding posttest. O₃ and O₄ indicate the pretest and posttest administered to the comparison group, respectively. X₁ refers to the implementation of Edu-Board Game-Assisted STAD in the experimental condition, whereas X₂ refers to Problem-Based Learning assisted by ice-cream-stick manipulatives in the comparison condition.

The intervention was administered at the class level. To support consistency of implementation, both conditions followed predetermined lesson plans with the same content scope, instructional duration, and pretest-posttest assessment framework. Implementation was observed using a lesson-implementation checklist to determine whether the planned instructional stages were carried out as intended. In the experimental condition, observation focused on the principal STAD stages, including class presentation, team organization and study, use of the Edu-Board Game, individual quizzes, calculation of improvement scores, and group recognition. The comparison condition was monitored against the planned PBL sequence. Observer records indicated that the instructional procedures were implemented without notable disruptions to content coverage, session structure, or instructional sequence.

Participants and Research Context

The study was conducted with Grade IV students at SD Negeri Mekar Mukti 06, Bekasi Regency, Indonesia. The target population comprised students enrolled in Grade IV during the semester in which the study was implemented. The research sample consisted of 60 students drawn from two intact classes, with 30 students assigned to the experimental condition and 30 students to the comparison condition. Participants were selected purposively based on the willingness of the school and teachers to participate and the availability of intact classes that could be used for the two instructional conditions. The students were between 9 and 10 years old and included 26 boys and 34 girls. The participating school was a public elementary school with an "A" accreditation status.

Eligibility for participation required students to be actively enrolled in Grade IV and to complete the full sequence of research activities, including the pretest, instructional sessions, and posttest. Recruitment was coordinated with the school principal and classroom teachers.

No students withdrew during the study, resulting in complete data for all 60 participants. The sample size was not determined through an a priori statistical power analysis; rather, it reflected the available intact classes and the practical constraints of the school setting. This should be considered when interpreting the precision of the statistical estimates and the study's ability to detect small between-group differences.

Because the two instructional conditions were implemented in separate intact classes, the class served as the practical unit of treatment delivery, although student-level scores were used for the statistical analyses. Each instructional condition was represented by only one class. This structure introduces the possibility that class-specific characteristics, including differences in classroom dynamics or teacher-related factors, may be associated with the observed outcomes. Accordingly, the study was designed to provide a comparative examination of the two instructional conditions within the participating school context rather than a broadly generalizable estimate of treatment effectiveness.

The research involved minors and was therefore conducted with attention to ethical requirements for educational research. Permission to conduct the study was obtained from the school principal, and the classroom teachers approved the instructional procedures used during the study. Informed consent was obtained from the parents or guardians of all participating students prior to data collection. Participation was voluntary, and students' data were treated confidentially and used only for research purposes.

Research Instrument

Students' mathematical problem-solving ability was measured using a constructed-response test developed for the topics of Least Common Multiple (LCM) and Greatest Common Factor (GCF). The instrument was structured around four stages of mathematical problem solving: understanding the problem, devising a plan, carrying out the plan, and checking the solution. These stages were adapted from Pólya's problem-solving framework and operationalized as observable indicators in students' written responses (Pólya, 1957; Riyanto et al., 2024). The test consisted of five contextual problems, with each item requiring students to demonstrate all four stages of the problem-solving process. The specification of the instrument is presented in Table 2.

Table 2. Mathematical Problem-Solving Test Specification

No.	Problem-Solving Indicator	Operational Description	Item Number	Content Domain
1	Understanding the Problem	Students identify relevant information and determine what is being asked in the problem.	1, 2, 3, 4, 5	LCM and GCF applications
2	Devising a Plan	Students formulate an appropriate solution strategy based on the information identified.	1, 2, 3, 4, 5	LCM and GCF applications
3	Carrying Out the Plan	Students apply the planned strategy using appropriate mathematical procedures and calculations.	1, 2, 3, 4, 5	LCM and GCF applications
4	Checking the Solution	Students review the solution process and formulate a conclusion based on the result.	1, 2, 3, 4, 5	LCM and GCF applications

Students' responses were scored analytically using a rubric aligned with the four problem-solving indicators. Understanding the problem and checking the solution were each scored from 0 to 2, whereas devising a plan and carrying out the plan were each scored from 0 to 3. Thus, each item had a maximum score of 10, producing a maximum raw score of 50 across the five items. Raw scores were subsequently converted to a 0-100 scale for statistical analysis. The detailed scoring criteria are presented in Table 3.

Table 3. Scoring Rubric for Mathematical Problem-Solving Ability

Indicator	Score	Scoring Criteria
Understanding the Problem	0	No relevant information or statement of what is being asked is provided.
	1	Relevant information and the question are identified, but the response is incomplete or contains inaccuracies.
	2	Relevant information and the question are identified accurately and completely.
Devising a Plan	0	No strategy or procedural plan for solving the problem is provided.
	1	A strategy is proposed, but it is inappropriate for the problem.
	2	The selected strategy is appropriate but incomplete.
	3	The selected strategy is appropriate and complete.
Carrying Out the Plan	0	No calculation or solution procedure is attempted.
	1	A solution attempt is made, but the mathematical procedure is unclear or substantially incorrect.
	2	The mathematical procedure is generally correct, but an error occurs in the calculation or final result.
	3	The mathematical procedure is carried out logically and the final result is correct.
Checking the Solution	0	No attempt is made to review the solution or formulate a conclusion.
	1	The solution is reviewed or a conclusion is provided, but it is incomplete or contains an inconsistency.
	2	The solution is appropriately reviewed and a valid conclusion is formulated.

Before its use in the main study, the instrument underwent expert review and empirical pilot testing. Content relevance and clarity were reviewed by two experts, consisting of a mathematics education lecturer and an elementary school teacher, using a five-point validation scale. Because the original study records reported the expert-review procedure but did not provide a separate quantitative content-validity coefficient, the expert review is treated here as evidence of content-based evaluation rather than as a numerical validity estimate. The revised instrument was then pilot-tested with 30 students who were not included in the main research sample.

The pilot data were analyzed using ANATES to examine item discrimination, difficulty level, item-total correlation, and test reliability. As shown in Table 4, discrimination indices ranged from 30.00% to 43.13%, and all five items were classified as having a medium level of difficulty. Item-total correlation coefficients ranged from 0.774 to 0.890, while the overall test reliability coefficient was 0.86. These results indicated that the five items demonstrated satisfactory empirical characteristics for use in the pretest and posttest assessments.

Table 4. Item Analysis Results for the Mathematical Problem-Solving Test

Item	Discrimination Index (%)	Difficulty Level	Item-Total Correlation	Correlation Significance
1	30.00	Medium	0.890	Very significant
2	32.50	Medium	0.860	Very significant
3	43.13	Medium	0.878	Very significant
4	33.13	Medium	0.774	Very significant
5	33.75	Medium	0.885	Very significant

Note. Pilot sample = 30 students; mean test score = 50.63; SD = 14.36; XY correlation = 0.76; test reliability = 0.86.

The same five-item instrument and scoring rubric were used for the pretest and posttest to maintain consistency in the measurement of mathematical problem-solving ability across the two measurement occasions. Responses were scored by a single rater. Consequently, inter-rater reliability could not be estimated, which should be considered when interpreting the measurement results. In addition, the instrument was specifically developed for LCM and GCF

applications; therefore, the scores represent mathematical problem-solving performance within this content domain rather than mathematical problem-solving ability across all elementary mathematics topics.

Intervention and Comparison Conditions

The experimental condition implemented Student Teams Achievement Divisions (STAD) supported by a physical educational board game entitled *LCM and GCF Adventure*. The board game was designed as a structured learning medium for practicing mathematical problem-solving on LCM and GCF. It consisted of a zigzag pathway with 54 spaces from START to FINISH, including designated LCM and GCF task spaces. Students played in small groups of three to five members and took turns rolling a die to move along the board. When a group landed on a designated task space, members were required to solve the corresponding mathematical problem collaboratively. A correct response allowed the group to remain in position and earned 10 points, whereas an incorrect response required the group to move two spaces backward. The overall design of the Edu-Board Game is shown in Figure 1.



Figure 1. Design of the Edu-Board Game

The game was supported by three sets of task cards consisting of 20 LCM question cards, 20 GCF question cards, and 20 Bonus cards. These cards were used to structure the mathematical tasks encountered during play and to provide repeated opportunities for students to discuss and solve problems within their teams. The question cards focused directly on LCM and GCF applications, while the Bonus cards provided additional game-based challenges within the learning activity. Examples of question and Bonus cards are presented in Figure 2.



Figure 2. Examples of Edu-Board Game Question and Bonus Cards

The experimental condition was implemented over three meetings, with each meeting lasting 2×35 minutes. The first meeting focused on LCM, the second on GCF, and the third on integrated LCM-GCF problem-solving activities. Each meeting followed the principal stages of STAD: class presentation, team study, individual assessment, calculation of individual improvement scores, and group recognition. The Edu-Board Game was incorporated primarily during the team-study stage to provide a structured context for collaborative problem solving. Thus, the experimental condition combined the cooperative structure of STAD with a physical, rule-based educational game rather than treating the board game as a separate activity outside the instructional model.

The comparison condition received Problem-Based Learning (PBL) supported by two-colored ice-cream-stick manipulatives. The sticks were used as concrete materials to help students explore multiples, factors, and related LCM and GCF relationships during problem-solving activities. The visual design of the manipulatives used in this condition is presented in Figure 3.

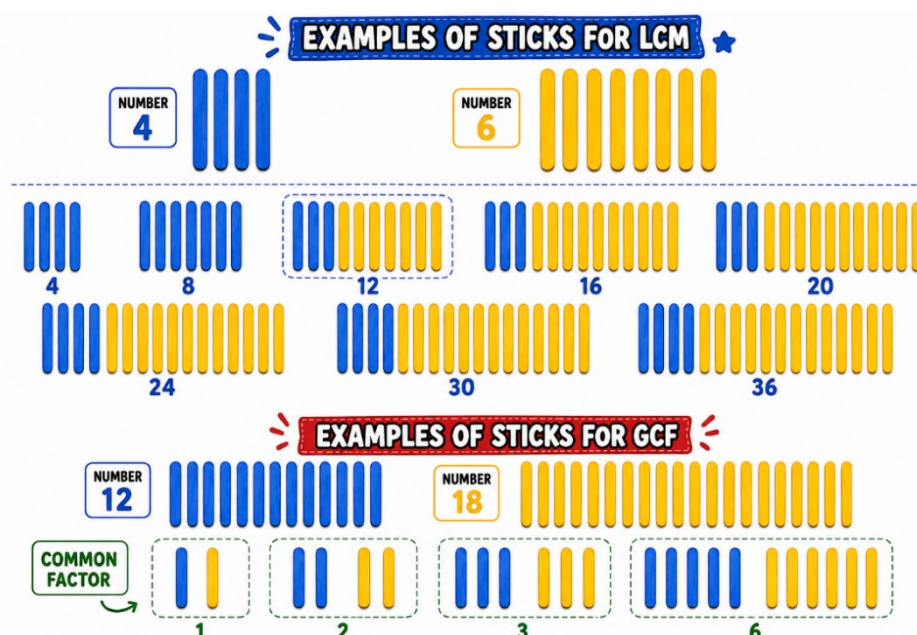


Figure 3. Ice-Cream-Stick Manipulatives Used in the PBL Condition

The PBL condition followed the stages of problem orientation, organization of students for learning, investigation using the ice-cream-stick manipulatives, presentation of results, and evaluation of the problem-solving process. Like the experimental condition, the comparison condition was conducted over three meetings with the same duration, content coverage, and assessment schedule. The principal distinction between the two instructional conditions therefore lay in the combination of pedagogical structure and learning medium: STAD integrated with a rule-based educational board game in the experimental condition, and PBL integrated with concrete manipulatives in the comparison condition.

Data Collection Procedure

Data collection was conducted in a sequence consisting of preparation, pretesting, implementation of the two instructional conditions, and posttesting. During the preparation stage, the researchers coordinated with the participating school, confirmed the two intact Grade IV classes, prepared the lesson plans and instructional media, and finalized the mathematical problem-solving test following expert review and pilot testing. The same assessment instrument was used for both measurement occasions to ensure consistency in evaluating students' mathematical problem-solving performance on LCM and GCF.

Before the instructional sessions began, all participating students completed the pretest to establish their initial level of mathematical problem-solving performance. The experimental class then participated in Edu-Board Game-Assisted STAD, whereas the comparison class received PBL assisted by ice-cream-stick manipulatives. Both instructional conditions were implemented over three meetings of 2×35 minutes, with equivalent content coverage, instructional duration, and assessment procedures. The first meeting focused on LCM, the second on GCF, and the third on integrated LCM-GCF problem-solving activities. After completion of the instructional sequence, all students completed the posttest using the same scoring framework as the pretest.

Implementation of both instructional conditions was monitored using a lesson-implementation checklist. An observer who was not involved in delivering the instruction recorded whether the planned instructional stages were implemented as specified. In the experimental condition, monitoring covered the main STAD stages, including class presentation, team study using the Edu-Board Game, individual assessment, calculation of improvement scores, and group recognition. The comparison condition was monitored against the planned PBL sequence. Observer records indicated that the instructional sessions were implemented according to the lesson plans, with no notable deviations in content coverage, session structure, or instructional sequence and no substantial disruptions such as absenteeism or scheduling problems.

Because the participants were elementary school students, ethical procedures were addressed before data collection. Permission to conduct the study was obtained from the school principal, instructional procedures were coordinated with the classroom teachers, and informed consent was obtained from the parents or guardians of all participating students. Participation covered the complete pretest, instructional, and posttest sequence, and no students withdrew during the study.

Data Analysis

Data were analyzed using SPSS version 26.0. Descriptive statistics were first calculated to summarize students' mathematical problem-solving performance in the two instructional conditions. Means, standard deviations, minimum and maximum scores, and 95% confidence intervals were reported for the pretest and posttest scores. Raw change scores were calculated as posttest minus pretest to describe the magnitude of improvement within each group. In addition, normalized gain (N-Gain) scores were calculated as a supplementary indicator of relative improvement from pretest to posttest. N-Gain values were interpreted using the conventional categories of low ($g < .30$), medium ($.30 \leq g < .70$), and high ($g \geq .70$) (Hake,

1999). These descriptive analyses were used primarily to address RQ1 by characterizing changes in mathematical problem-solving performance across the two instructional conditions.

Before conducting between-group analyses, the distributional assumptions of the data were examined. The Shapiro-Wilk test was used to assess normality because each group consisted of 30 students. Homogeneity of variance was examined using Levene's test, and linearity between pretest and posttest scores was assessed before baseline-adjusted analysis. Baseline comparability between the experimental and comparison groups was examined using an independent-samples *t*-test on pretest scores. Because the pretest variances were unequal, the baseline comparison was interpreted using the unequal-variances estimate where appropriate. Statistical significance was evaluated at $\alpha = .05$, and 95% confidence intervals were reported for the principal estimates.

To address RQ2, a baseline-adjusted multiple linear regression model was used, with posttest mathematical problem-solving score as the dependent variable and pretest score and instructional condition as predictors. Group membership was dummy-coded to estimate the adjusted difference between students receiving Edu-Board Game-Assisted STAD and those receiving ice-cream-stick-assisted PBL after accounting for initial mathematical problem-solving ability. The group coefficient therefore represented the estimated between-condition difference in posttest performance at a common level of baseline ability. The analysis reported the unstandardized regression coefficient (*B*), standard error, 95% confidence interval, test statistic, and *p* value. Adjusted posttest means were also derived from the baseline-adjusted model to facilitate interpretation of the between-group comparison.

N-Gain scores were retained as a supplementary analysis rather than as the primary test of between-group effectiveness. Mean N-Gain scores were compared between the two conditions using an independent-samples *t*-test after checking normality and homogeneity of variance. Cohen's *d*, together with its 95% confidence interval, was reported to describe the standardized magnitude of the N-Gain difference. Because the groups differed substantially at baseline, the N-Gain comparison was interpreted cautiously and did not replace the baseline-adjusted model as the principal analysis for evaluating between-condition differences.

The analysis strategy therefore distinguished between two forms of evidence. Pretest, posttest, raw change, and N-Gain statistics were used to describe how performance changed within each instructional condition, whereas the baseline-adjusted regression model was used to evaluate whether posttest mathematical problem-solving performance differed between the two conditions after accounting for students' initial ability. This distinction was necessary because the nonequivalent-group design produced a significant pretest difference between the two intact classes, making an unadjusted posttest comparison insufficient for interpreting between-group differences.

RESULTS AND DISCUSSION

Descriptive and Baseline Results

Descriptive statistics for students' mathematical problem-solving performance are presented in Table 5. Both instructional conditions showed substantial increases in mean scores from pretest to posttest. In the comparison group, the mean score increased from 26.27 (SD = 8.03) at pretest to 60.53 (SD = 9.35) at posttest, corresponding to a mean raw change of 34.27 points (SD = 9.86). In the experimental group, the mean score increased from 41.33 (SD = 15.34) at pretest to 73.67 (SD = 11.48), with a mean raw change of 32.33 points (SD = 7.74). Although both groups improved, the experimental group entered the intervention with a considerably higher level of mathematical problem-solving performance.

Table 5. Descriptive Statistics for Mathematical Problem-Solving Performance

Group	Measure	N	Mean	SD	95% CI	Min.	Max.
Comparison	Pretest	30	26.27	8.03	[23.27, 29.26]	13	40
	Posttest	30	60.53	9.35	[57.04, 64.02]	40	78

Group	Measure	N	Mean	SD	95% CI	Min.	Max.
Experimental	Raw change	30	34.27	9.86	[30.58, 37.95]	11	48
	Pretest	30	41.33	15.34	[35.60, 47.06]	14	82
	Posttest	30	73.67	11.48	[69.38, 77.95]	48	100
	Raw change	30	32.33	7.74	[29.44, 35.22]	14	46

Note. Raw change was calculated as posttest minus pretest.

The two groups differed significantly in their initial mathematical problem-solving performance. The experimental group had a substantially higher pretest mean than the comparison group, $t(43.76) = -4.77$, $p < .001$. This baseline difference indicated that direct comparison of the posttest scores alone would not adequately account for differences already present before implementation of the instructional conditions. Accordingly, baseline performance was incorporated into the principal between-group analysis.

Assumption Testing

The assumptions underlying the statistical analyses were examined before inferential testing. Results of the Shapiro-Wilk tests indicated no statistically significant departures from normality for the pretest or posttest scores in either instructional condition, with all p values greater than .05, as presented in Table 6.

Table 6. Shapiro-Wilk Normality Test

Group and Measure	Statistic	df	p
Comparison pretest	.962	30	.355
Comparison posttest	.976	30	.711
Experimental pretest	.949	30	.157
Experimental posttest	.984	30	.924

Levene's test indicated that the pretest variances differed significantly between the two groups, $F(1, 58) = 7.338$, $p = .009$. In contrast, the posttest variances did not differ significantly, $F(1, 58) = 0.906$, $p = .345$. The results are presented in Table 7.

Table 7. Levene's Test of Homogeneity of Variance

Measure	Levene Statistic	df1	df2	p
Pretest	7.338	1	58	.009
Posttest	0.906	1	58	.345

The unequal pretest variances, together with the significant baseline difference between groups, reinforced the need to account for initial mathematical problem-solving performance when evaluating between-group posttest outcomes.

Changes in Mathematical Problem-Solving Performance

To address RQ1, changes in mathematical problem-solving performance from pretest to posttest were examined within each instructional condition. Both groups showed substantial improvement. The comparison group increased by an average of 34.27 raw-score points, whereas the experimental group increased by 32.33 points. The experimental group nevertheless retained the higher absolute performance level at posttest because it had also started from a considerably higher baseline. Normalized gain was additionally calculated to describe improvement relative to the maximum possible score. As shown in Table 8, the experimental group obtained a mean N-Gain of 0.57, corresponding to 57.12%, whereas the comparison group obtained a mean N-Gain of 0.46, corresponding to 46.34%. Both values were classified within the medium-gain category.

Table 8. Normalized Gain in Mathematical Problem-Solving Performance

Group	Mean N-Gain	N-Gain (%)	Category
Experimental	0.5712	57.12	Medium
Comparison	0.4634	46.34	Medium

These findings indicate that students in both instructional conditions improved their mathematical problem-solving performance during the study period. The experimental condition produced the higher normalized gain, although both groups remained within the same medium-gain category.

Baseline-Adjusted Between-Group Comparison

To address RQ2, posttest mathematical problem-solving performance was compared between the two instructional conditions after accounting for students' pretest scores. The baseline-adjusted analysis indicated that the estimated difference favored Edu-Board Game-Assisted STAD. However, the group effect did not reach the conventional threshold for statistical significance, $B = 4.04$, 95% CI $[-0.53, 8.60]$, $F(1, 57) = 3.14$, $p = .082$, partial eta squared (η^2) = .052.

The adjusted posttest mean was 68.75 ($SE = 1.52$) for the experimental group and 63.73 ($SE = 1.88$) for the comparison group. The baseline-adjusted model results and adjusted group means are presented together in Table 9.

Table 9. Baseline-Adjusted Comparison of Posttest Mathematical Problem-Solving Performance

Result	Estimate	SE	95% CI	F	df	p	η^2
Instructional condition	4.04	—	$[-0.53, 8.60]$	3.14	1, 57	.082	.052
Experimental adjusted mean	68.75	1.52	—	—	—	—	—
Comparison adjusted mean	63.73	1.88	—	—	—	—	—

Note. Adjusted means represent estimated posttest scores after accounting for pretest mathematical problem-solving performance.

The confidence interval for the instructional-condition coefficient included zero, and the p value exceeded .05. Therefore, although adjusted posttest performance numerically favored Edu-Board Game-Assisted STAD, the present data did not provide statistically conclusive evidence of a difference between the two instructional conditions after baseline mathematical problem-solving performance was taken into account.

Supplementary N-Gain Analysis

A supplementary between-group analysis was conducted using N-Gain percentages. The experimental group achieved a higher mean N-Gain percentage ($M = 57.12$, $SD = 12.81$) than the comparison group ($M = 46.34$, $SD = 12.51$). The N-Gain distributions met the normality assumption in both groups, with $p = .058$ for the experimental group and $p = .141$ for the comparison group. Levene's test also indicated homogeneous variances, $F = 0.068$, $p = .795$. The independent-samples t -test indicated a statistically significant difference in N-Gain percentages between the two instructional conditions, $t(58) = 3.30$, $p = .002$, with a mean difference of 10.78 percentage points. The results are presented in Table 10.

Table 10. Comparison of N-Gain Percentages Between Instructional Conditions

Group	N	Mean	SD
Experimental	30	57.12	12.81
Comparison	30	46.34	12.51

Note. Levene's $F = 0.068$, $p = .795$; $t(58) = 3.30$, $p = .002$; mean difference = 10.78 percentage points.

The standardized difference in N-Gain was large, with Cohen's $d = 0.851$ and a 95% CI $[0.319, 1.377]$. Similar estimates were obtained using Hedges' correction and Glass's delta, as shown in Table 11.

Table 11. Effect Size Estimates for the N-Gain Difference

Effect-Size Estimator	Estimate	95% CI
Cohen's d	0.851	$[0.319, 1.377]$
Hedges' correction	0.840	$[0.315, 1.359]$
Glass's delta	0.862	$[0.303, 1.407]$

The supplementary N-Gain analysis therefore favored the experimental condition. However, this finding should be interpreted in conjunction with the baseline-adjusted analysis because the two intact classes differed substantially before the intervention. Accordingly, the N-Gain result provides complementary evidence of relative improvement, whereas the baseline-adjusted analysis provides the principal evidence for evaluating RQ2.

Discussion

The findings indicate that both instructional conditions were associated with substantial improvement in students' mathematical problem-solving performance. The experimental group achieved a higher normalized gain than the comparison group, suggesting greater relative improvement across the study period. However, after students' initial mathematical problem-solving ability was taken into account, the adjusted posttest difference favored Edu-Board Game-Assisted STAD numerically but did not reach statistical significance. This pattern suggests that the experimental condition supported meaningful learning progress, but the available evidence does not demonstrate a statistically reliable advantage over the PBL comparison. Such a result is consistent with broader evidence showing that instructional innovations in elementary mathematics may produce positive learning gains without necessarily outperforming other well-designed active approaches. Reviews of mathematics interventions likewise indicate that observed effects vary substantially according to instructional design, comparison condition, learner characteristics, and implementation context (Bozkuş et al., 2026; Kyndt et al., 2013; Pellegrini et al., 2021).

The improvement observed in both groups is pedagogically plausible because the comparison condition was itself an active form of instruction rather than a passive or lecture-based control. Students in the comparison group participated in PBL supported by concrete ice-cream-stick manipulatives, which required them to investigate mathematical relationships, develop solutions, and engage directly with LCM and GCF problems. Research on problem-based approaches in mathematics indicates that such instruction can promote mathematical reasoning and problem solving by organizing learning around investigation, solution development, explanation, and reflection (Aba-Oli et al., 2025; Ayari et al., 2025). The use of concrete manipulatives may further support mathematical understanding when physical objects are explicitly connected to the concepts they represent. Meta-analytic and cognitive evidence indicates that manipulatives can benefit mathematics learning, although their effectiveness depends on how clearly learners are guided from concrete representations toward underlying mathematical structures (Carbonneau et al., 2013; Fyfe et al., 2014; Laski et al., 2015; Pouw et al., 2014). The substantial improvement observed in the comparison group therefore reinforces the interpretation that both instructional conditions provided meaningful opportunities for mathematical learning.

The performance pattern in the experimental group can be interpreted in relation to the complementary roles of cooperative learning and rule-based game activity. STAD structures learning through team study while maintaining individual responsibility, whereas the Edu-Board Game embeds mathematical problems within a sequence of moves, questions, feedback, and task consequences. Research on cooperative learning has shown that structured peer interaction can support academic achievement when group tasks are clearly organized and individual accountability is maintained (Kyndt et al., 2013; Pellegrini et al., 2021). Game-based learning may add repeated opportunities for retrieval, decision making, feedback, and interaction around disciplinary content. Large-scale reviews suggest that games can improve cognitive learning outcomes, but their educational value depends less on the mere presence of a game format than on the instructional features incorporated into it, including feedback, repeated practice, alignment with learning objectives, and integration with classroom pedagogy (Clark et al., 2016; Wouters et al., 2013). In the present study, the board game was embedded within the STAD team-study phase rather than functioning as an isolated recreational activity,

which provides a plausible pedagogical basis for the improvement observed in the experimental group.

The findings also align with a broader body of research showing that the effects of game-based mathematics instruction are generally positive but heterogeneous. Meta-analytic evidence across K–12 mathematics has reported benefits of game-based learning for mathematical performance, while also demonstrating considerable variation according to game design, duration, targeted outcome, and instructional comparison (Bozkuş et al., 2026; Byun & Joung, 2018). Similar conclusions have been reported for higher-order mathematical thinking, where game-based approaches produced favorable effects but with substantial variability across studies (Anggoro et al., 2025). Affective outcomes show a similarly differentiated pattern. Game-based interventions can support mathematics motivation and reduce anxiety in some contexts, but their effects depend on the type of game, duration of exposure, age of learners, and the specific outcome being assessed (Dondio et al., 2023; Fadda et al., 2022). These findings support a cautious interpretation of the present results: integrating a physical board game with STAD may provide a productive learning environment, but the presence of game features alone should not be expected to produce a consistently large advantage over another active instructional approach.

The lack of a statistically conclusive adjusted difference is also understandable when the nature of the comparison condition is considered. When an intervention is compared with relatively passive instruction, the estimated advantage may reflect not only the intervention itself but also a broader contrast in the amount and quality of student activity. In the present study, however, both groups engaged in structured problem solving, interaction with instructional materials, and teacher-guided learning activities. Evidence from research syntheses suggests that the added value of a new instructional approach often becomes smaller when the comparison group also receives purposeful and active instruction (Benavides-Varela et al., 2020; Topping et al., 2022). For example, digital mathematics interventions may generate positive learning outcomes without consistently outperforming other structured digital or problem-centered approaches (Benavides-Varela et al., 2020). Likewise, systematic evidence on PBL shows that problem-centered instruction itself can support mathematical reasoning and achievement (Aba-Oli et al., 2025; Ayari et al., 2025). The nonsignificant adjusted difference in the present study should therefore not be interpreted as evidence that Edu-Board Game-Assisted STAD lacked educational value, but rather that its additional advantage over an already active comparison condition was not statistically established.

Another relevant feature of the findings is that both instructional conditions incorporated forms of concrete or representational activity. Concepts such as factors, multiples, LCM, and GCF require students to coordinate numerical relationships that may remain abstract when taught only symbolically. Research has long shown that concrete materials can facilitate mathematical understanding when learners recognize the relationship between the physical representation and the mathematical concept it represents (Uttal et al., 1997). More recent evidence emphasizes that effective mathematics learning requires students to move beyond manipulation itself and connect concrete actions to generalized mathematical relations and symbolic representations (Fyfe et al., 2014; Pouw et al., 2014). In the comparison condition, the ice-cream sticks represented factor and multiple relationships, whereas in the experimental condition mathematical reasoning was embedded within question cards and movement through a rule-governed board. Both approaches may therefore have supported students in linking concrete activity with mathematical reasoning through different representational pathways. Research on strategy comparison and mathematical cognition further suggests that students benefit when they are encouraged to interpret, compare, and justify solution approaches rather than merely reproduce a single procedure (Rittle-Johnson et al., 2020). Feedback embedded within game-based environments may also contribute to this process when it is directly connected to students' mathematical responses and misconceptions (Yang & Lu, 2021).

The significant N-Gain difference and the nonsignificant baseline-adjusted result should be understood as providing different, though related, information about students' performance. N-Gain expresses improvement relative to the maximum score available above a student's initial score and is therefore useful for describing relative progress. However, normalized gain can be associated with differences in pretest distributions, and its interpretation in nonequivalent groups remains methodologically debated. Some researchers have shown that normalized gain may exhibit relationships with initial scores and have therefore recommended caution when it is used for between-group comparisons (Nissen et al., 2018). Others have argued that normalized gain remains informative when student ability and the mathematical properties of the measure are properly considered (Coletta & Steinert, 2020). More generally, analyses of pretest-posttest designs favor adjustment for baseline scores when the principal question concerns whether groups differ at posttest after accounting for their initial performance (Vickers & Altman, 2001). The two findings should therefore be viewed as complementary rather than contradictory. The N-Gain analysis indicates that the experimental group achieved greater relative improvement, whereas the baseline-adjusted analysis addresses more directly whether posttest performance differed between the two instructional conditions at a comparable level of initial ability.

Taken together, these findings contribute to a more differentiated understanding of game-supported cooperative learning in elementary mathematics. Edu-Board Game-Assisted STAD was associated with substantial improvement in mathematical problem-solving performance and produced a higher normalized gain than the PBL comparison. At the same time, the absence of a statistically significant adjusted posttest difference indicates that positive change within an instructional condition does not necessarily translate into clear superiority over another active approach. This interpretation is consistent with evidence that mathematics learning outcomes are shaped by the interaction among pedagogical structure, representations, feedback, learner characteristics, and the nature of the comparison condition rather than by the nominal instructional model alone (Bozkuş et al., 2026; Clark et al., 2016; Kyndt et al., 2013). The principal contribution of the present study therefore lies in showing that integrating a physical educational board game within STAD represents a viable active learning configuration for elementary mathematical problem solving, while also demonstrating that its relative advantage must be interpreted in relation to the strength of the instructional alternative against which it is compared.

CONCLUSION

This study examined changes in Grade IV students' mathematical problem-solving performance under two active instructional conditions: Edu-Board Game-Assisted STAD and Problem-Based Learning supported by ice-cream-stick manipulatives. Both groups showed substantial improvement from pretest to posttest, indicating that each instructional condition provided meaningful opportunities for students to develop problem-solving performance in LCM and GCF. The experimental group achieved the higher normalized gain, suggesting greater relative improvement during the intervention period.

When initial mathematical problem-solving ability was taken into account, however, the adjusted posttest difference between the two instructional conditions was not statistically significant. Although the adjusted mean remained numerically higher for the Edu-Board Game-Assisted STAD group, the evidence was insufficient to establish a reliable posttest advantage over the active PBL comparison. The significant N-Gain difference should therefore be interpreted as complementary evidence of relative improvement rather than as definitive evidence of the superiority of the experimental condition.

The findings indicate that integrating a physical, rule-based educational board game within STAD represents a viable instructional approach for elementary mathematical problem solving. Its value lies in combining structured cooperative learning with repeated mathematical tasks embedded in an interactive game environment. At the same time, the results demonstrate

the importance of evaluating instructional innovations against meaningful active alternatives and of considering students' initial performance when interpreting comparative learning outcomes.

LIMITATION

The findings should be interpreted within several methodological constraints. The study involved only two intact classes from a single elementary school, with one class representing each instructional condition and no random assignment at the individual level. Because different teachers were involved across the two classes, instructional-condition effects could not be fully separated from teacher- or class-specific influences. The intervention was also relatively short, consisting of three meetings, and therefore captured immediate post-intervention performance rather than longer-term retention or transfer. Mathematical problem-solving was assessed using five constructed-response items restricted to LCM and GCF, and responses were scored by a single rater, preventing estimation of inter-rater reliability. In addition, the study did not directly assess potential process variables such as collaboration, engagement, motivation, or communication, so these mechanisms cannot be treated as empirically established explanations for the observed performance patterns.

RECOMMENDATION

Future research should examine Edu-Board Game-Assisted STAD across multiple classes and schools using randomized, matched, or cluster-randomized designs where feasible, with sufficient class-level replication to distinguish instructional effects from teacher and classroom influences. Longer interventions and delayed posttests are also recommended to determine whether observed improvements are sustained and transferable to other mathematical topics. Subsequent studies should employ broader problem-solving measures, involve multiple raters for constructed-response scoring, and include validated measures of collaboration, engagement, motivation, and other learning processes that may help explain how cooperative board-game-based instruction supports mathematical problem solving.

ACKNOWLEDGMENTS

Deep appreciation is dedicated to all institutions and elementary school colleagues, including principals, teachers, and students, for their full participation and support during the data collection process. The author also appreciates the supervisors and various related parties who constantly provide direction, constructive input, and assistance until this manuscript is completed. All of these contributions have a crucial role in the success of this research.

FUNDING INFORMATION

This research received no external funding. The article processing charge (APC) was funded by the authors.

AUTHOR CONTRIBUTIONS STATEMENT

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Febby Deca Lestari	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	
Hafiziani Eka Putri	✓	✓		✓	✓			✓		✓	✓	✓		
Teten Ginanjar Rahayu	✓	✓		✓	✓			✓		✓	✓	✓		
Alfiana Nurussama												✓		
Mariano Dos Santos										✓		✓		

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

ETHICAL APPROVAL

Permission to conduct the study was obtained from the participating school. The research procedures were conducted in accordance with applicable institutional policies, and students' data were treated confidentially and used solely for research purposes.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request. As the data involve minors, access may be subject to ethical and privacy considerations

REFERENCES

- Aba-Oli, Z., Koyas, K., & Husen, A. (2025). Higher-order thinking skills-oriented problem-based learning interventions in mathematics: A systematic literature review. *School Science and Mathematics*, 125(3), 214–231. <https://doi.org/10.1111/ssm.12676>
- Anggoro, B. S., Dewantara, A. H., Suherman, S., Muhammad, R. R., & Saraswati, S. (2025). Effect of game-based learning on students' mathematics high order thinking skills: A meta-analysis. *Revista de Psicodidáctica*, 30(1), 500158. <https://doi.org/10.1016/j.psicoe.2024.500158>
- Asmedy, A. (2021). Perbandingan hasil belajar matematika siswa yang diajar dengan model pembelajaran posing problem berkelompok dan metode ceramah. *Jurnal Inovasi, Evaluasi dan Pengembangan Pembelajaran*, 1(2), 69–75. <https://doi.org/10.54371/jiepp.v1i2.107>
- Ayari, M. A., Sellami, A., Santhosh, M. E., Naji, K. K., Al-Ali, A., & Al-Hazbi, S. M. A. (2025). From problems to performance: A systematic review of problem-based learning in K–12 mathematics. *Frontiers in Education*, 10, 1731307. <https://doi.org/10.3389/educ.2025.1731307>
- Azzahra, P. L., Rusli, R. K., & Dhanial, H. (2026). Pengaruh media pembelajaran board game terhadap hasil belajar siswa kelas V sekolah dasar. *Al-Kaff: Jurnal Sosial Humaniora*, 4(1), 89–95. <https://doi.org/10.30997/al-kaff.v4i1.20898>
- Barreto, D., Oyarzun, B., & Conklin, S. (2022). Integration of cooperative learning strategies in online settings. *E-Learning and Digital Media*, 19(6), 574–594. <https://doi.org/10.1177/20427530221104187>
- Benavides-Varela, S., Zandonella Callegher, C., Fagiolini, B., Leo, I., Altoè, G., & Lucangeli, D. (2020). Effectiveness of digital-based interventions for children with mathematical learning difficulties: A meta-analysis. *Computers & Education*, 157, 103953. <https://doi.org/10.1016/j.compedu.2020.103953>
- Bozkuş, F., Kaya, M., & Arslan, Z. (2026). Effects of game-based learning on student outcomes in mathematics education: A second-order meta-analysis. *Educational Research Review*, 52, 100816. <https://doi.org/10.1016/j.edurev.2026.100816>
- Bramantha, H., & Meliandani, R. (2024). Perbedaan hasil belajar antara penggunaan model pembelajaran Course Review Horay (CRH) dengan metode ceramah pada siswa kelas III sekolah dasar. *Mutiara PGSD*, 1(1), 1–10. <https://doi.org/10.63440/mutiarapgsd.v1i1.5>
- Byun, J., & Joung, E. (2018). Digital game-based learning for K–12 mathematics education: A meta-analysis. *School Science and Mathematics*, 118(3–4), 113–126. <https://doi.org/10.1111/ssm.12271>
- Carbonneau, K. J., Marley, S. C., & Selig, J. P. (2013). A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. *Journal of Educational Psychology*, 105(2), 380–400. <https://doi.org/10.1037/a0031084>
- Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, 86(1), 79–122. <https://doi.org/10.3102/0034654315582065>
- Coletta, V. P., & Steinert, J. J. (2020). Why normalized gain should continue to be used in analyzing preinstruction and postinstruction scores on concept inventories. *Physical Review Physics Education Research*, 16(1), 010108. <https://doi.org/10.1103/PhysRevPhysEducRes.16.010108>
- Damayanti, A., Arismundar, & Tolla, I. (2025). Pengaruh penggunaan media pembelajaran interaktif terhadap minat belajar siswa. *Jurnal Paris Langkis*, 6(1), 62–68. <https://doi.org/10.37304/jjpl.v6i1.17270>
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining “gamification.” In *Proceedings of the 15th International*

- Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9–15). Association for Computing Machinery. <https://doi.org/10.1145/2181037.2181040>
- Dondio, P., Gusev, V., & Rocha, M. (2023). Do games reduce maths anxiety? A meta-analysis. *Computers & Education*, 194, 104650. <https://doi.org/10.1016/j.compedu.2022.104650>
- Drestanta, I. N., Nelvita, A. W., Priskasari, A. D., Latifa, A.-N. K., & Novianto, A. (2025). Pengembangan media board game VOKUBA (Volume Kubus dan Balok) pada pembelajaran matematika kelas V sekolah dasar. *Edukasi: Jurnal Penelitian dan Artikel Pendidikan*, 17(1), 577–598. <https://doi.org/10.31603/edukasi.v17i1.13508>
- Fadda, D., Pellegrini, M., Vivanet, G., & Zandonella Callegher, C. (2022). Effects of digital games on student motivation in mathematics: A meta-analysis in K–12. *Journal of Computer Assisted Learning*, 38(1), 304–325. <https://doi.org/10.1111/jcal.12618>
- Fathurrohman, M., Nindiasari, H., & Rahayu, I. (2022). A conventional and digital mathematical board game design and development for use by students in learning arithmetic. *Journal on Mathematics Education*, 13(4), 631–660. <https://doi.org/10.22342/jme.v13i4.pp631-660>
- Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness fading in mathematics and science instruction: A systematic review. *Educational Psychology Review*, 26(1), 9–25. <https://doi.org/10.1007/s10648-014-9249-3>
- Hake, R. R. (1999). *Analyzing change/gain scores*. Indiana University.
- Huang, J., Cai, Y., Lv, Z., Huang, Y., & Zheng, X.-L. (2024). Toward self-regulated learning: Effects of different types of data-driven feedback on pupils' mathematics word problem-solving performance. *Frontiers in Psychology*, 15, 1356852. <https://doi.org/10.3389/fpsyg.2024.1356852>
- Kusmiyantoro, O. D., Sulianto, J., & Purnamasari, I. (2026). Studi komparasi model Problem Based Learning dan Discovery Learning ditinjau dari kemampuan pemecahan masalah matematika siswa SD N 1 Sulursari. *JP3: Jurnal Pendidikan dan Profesi Pendidik*, 12(1), 101–109. <https://doi.org/10.26877/jp3.v12i1.547>
- Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., & Dochy, F. (2013). A meta-analysis of the effects of face-to-face cooperative learning: Do recent studies falsify or verify earlier findings? *Educational Research Review*, 10, 133–149. <https://doi.org/10.1016/j.edurev.2013.02.002>
- Laski, E. V., Jor'dan, J. R., Daoust, C., & Murray, A. K. (2015). What makes mathematics manipulatives effective? Lessons from cognitive science and Montessori education. *SAGE Open*, 5(2), 1–8. <https://doi.org/10.1177/2158244015589588>
- Lin, Y.-T., & Cheng, C.-T. (2022). Effects of technology-enhanced board game in primary mathematics education on students' learning performance. *Applied Sciences*, 12(22), 11356. <https://doi.org/10.3390/app122211356>
- Nguyen, T. N. T., & Oanh, D. T. K. (2025). Cooperative learning and its influences on student engagement. *Cogent Education*, 12(1), 2513414. <https://doi.org/10.1080/2331186X.2025.2513414>
- Nissen, J. M., Talbot, R. M., Thompson, A. N., & Van Dusen, B. (2018). Comparison of normalized gain and Cohen's d for analyzing gains on concept inventories. *Physical Review Physics Education Research*, 14(1), 010115. <https://doi.org/10.1103/PhysRevPhysEducRes.14.010115>
- Nurarini, A. G., & Mariana, N. (2024). Analisis kesulitan peserta didik berkemampuan matematis tinggi dalam menyelesaikan soal numerasi bilangan pada Asesmen Kompetensi Minimum (AKM) kelas V sekolah dasar. *Jurnal Penelitian Pendidikan Guru Sekolah Dasar*, 12(7), 1352–1366.
- Nurmeliyanti, A., Nuraeni, R., & Puspitasari, N. (2024). Kemampuan pemecahan masalah matematis siswa yang mendapatkan pembelajaran Student Teams Achievement Division

- dan Problem Based Learning. *Plusminus: Jurnal Pendidikan Matematika*, 4(2), 385–396. <https://doi.org/10.31980/plusminus.v4i2.2223>
- Pellegrini, M., Lake, C., Neitzel, A., & Slavin, R. E. (2021). Effective programs in elementary mathematics: A meta-analysis. *AERA Open*, 7, 1–29. <https://doi.org/10.1177/2332858420986211>
- Pouw, W. T. J. L., van Gog, T., & Paas, F. (2014). An embedded and embodied cognition review of instructional manipulatives. *Educational Psychology Review*, 26(1), 51–72. <https://doi.org/10.1007/s10648-014-9255-5>
- Prang, M. Y., Salajang, S. M., & Mairing, J. P. (2024). Pengaruh pendekatan Pendidikan Matematika Realistik Indonesia terhadap kemampuan pemecahan masalah matematis siswa. *Kognitif: Jurnal Riset HOTS Pendidikan Matematika*, 4(3). <https://doi.org/10.51574/kognitif.v4i3.2033>
- Pólya, G. (1957). *How to solve it: A new aspect of mathematical method* (2nd ed.). Princeton University Press.
- Rittle-Johnson, B., Star, J. R., & Durkin, K. (2020). How can cognitive-science research help improve education? The case of comparing multiple strategies to improve mathematics learning and teaching. *Current Directions in Psychological Science*, 29(6), 599–609. <https://doi.org/10.1177/0963721420969365>
- Riyanto, O. R., Widyastuti, Yustitia, V., Oktaviyanthi, R., Sari, N. H. M., Izzati, N., Sukmaangara, B., Indartiningsih, D., Wibowo, A., Maharbid, D. A., & Wahid, S. (2024). *Kemampuan matematis*. CV. Zenius Publisher.
- Roorda, G., de Vries, S., & Smale-Jacobse, A. E. (2024). Using lesson study to help mathematics teachers enhance students' problem-solving skills with teaching through problem solving. *Frontiers in Education*, 9, 1331674. <https://doi.org/10.3389/educ.2024.1331674>
- Safitri, W. C. D. (2019). Efektivitas media board game terhadap kemampuan pemecahan masalah pada pembelajaran tematik di SD. *Mimbar PGSD Undiksha*, 7(2), 72–78. <https://doi.org/10.23887/jjpgsd.v7i2.17477>
- Sagita, D. K., Ermawati, D., & Riswari, L. A. (2023). Kemampuan pemecahan masalah matematis siswa sekolah dasar. *Jurnal Educatio FKIP UNMA*, 9(2), 431–439. <https://doi.org/10.31949/educatio.v9i2.4609>
- Santos-Trigo, M. (2024). Problem solving in mathematics education: Tracing its foundations and current research-practice trends. *ZDM–Mathematics Education*, 56(2), 211–222. <https://doi.org/10.1007/s11858-024-01578-8>
- Säfsström, A. I., Lithner, J., Palm, T., Palmberg, B., Sidenvall, J., Andersson, C., Boström, E., & Granberg, C. (2024). Developing a diagnostic framework for primary and secondary students' reasoning difficulties during mathematical problem solving. *Educational Studies in Mathematics*, 115(2), 125–149. <https://doi.org/10.1007/s10649-023-10278-1>
- Siregar, T. J. (2021). Peningkatan kemampuan pemecahan masalah matematis dan keterampilan sosial siswa SMP melalui pembelajaran kooperatif tipe STAD. *AXIOM: Jurnal Pendidikan dan Matematika*, 10(1), 97–109. <https://doi.org/10.30821/axiom.v10i1.9265>
- Siswanto, E., & Meiliasari, M. (2024). Kemampuan pemecahan masalah pada pembelajaran matematika: Systematic literature review. *Jurnal Riset Pembelajaran Matematika Sekolah*, 8(1), 45–59. <https://doi.org/10.21009/jrpms.081.06>
- Stewart, E., & Ball, L. (2023). The tension between allowing student struggle and providing support when teaching problem-solving in primary school mathematics. *Canadian Journal of Science, Mathematics and Technology Education*, 23, 791–817. <https://doi.org/10.1007/s42330-024-00309-1>
- Strohmaier, A. R., Reinhold, F., Hofer, S., Berkowitz, M., Vogel-Heuser, B., & Reiss, K. (2022). Different complex word problems require different combinations of cognitive

- skills. *Educational Studies in Mathematics*, 109(1), 89–114. <https://doi.org/10.1007/s10649-021-10079-4>
- Topping, K. J., Douglas, W., Robertson, D., & Ferguson, N. (2022). Effectiveness of online and blended learning from schools: A systematic review. *Review of Education*, 10(2), e3353. <https://doi.org/10.1002/rev3.3353>
- Uttal, D. H., Scudder, K. V., & DeLoache, J. S. (1997). Manipulatives as symbols: A new perspective on the use of concrete objects to teach mathematics. *Journal of Applied Developmental Psychology*, 18(1), 37–54. [https://doi.org/10.1016/S0193-3973\(97\)90013-7](https://doi.org/10.1016/S0193-3973(97)90013-7)
- Vessonen, T., Dahlberg, M., Hellstrand, H., Widlund, A., Korhonen, J., Aunio, P., & Laine, A. (2024). Task characteristics associated with mathematical word problem-solving performance among elementary school-aged children: A systematic review and meta-analysis. *Educational Psychology Review*, 36, 117. <https://doi.org/10.1007/s10648-024-09954-2>
- Vicente, S., Verschaffel, L., Sánchez, R., & Múñez, D. (2022). Arithmetic word problem solving: Analysis of Singaporean and Spanish textbooks. *Educational Studies in Mathematics*, 111(3), 375–397. <https://doi.org/10.1007/s10649-022-10169-x>
- Vickers, A. J., & Altman, D. G. (2001). Analysing controlled trials with baseline and follow up measurements. *BMJ*, 323(7321), 1123–1124. <https://doi.org/10.1136/bmj.323.7321.1123>
- Vita-Barrull, N., Estrada-Plana, V., March-Llanes, J., Sotoca-Orgaz, P., Guzmán, N., Ayesa, R., & Moya-Higueras, J. (2024). Do you play in class? Board games to promote cognitive and educational development in primary school: A cluster randomized controlled trial. *Learning and Instruction*, 93, 101946. <https://doi.org/10.1016/j.learninstruc.2024.101946>
- Wisenöcker, A. S., Binder, S., Holzer, M., Valentini, A., Wally, C., & Große, C. S. (2024). Mathematical problems in and out of school: The impact of considering mathematical operations and reality on real-life solutions. *European Journal of Psychology of Education*, 39, 767–783. <https://doi.org/10.1007/s10212-023-00718-0>
- Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, 105(2), 249–265. <https://doi.org/10.1037/a0031311>
- Yang, K.-H., & Lu, B.-C. (2021). Towards the successful game-based learning: Detection and feedback to misconceptions is the key. *Computers & Education*, 160, 104033. <https://doi.org/10.1016/j.compedu.2020.104033>
- Žakelj, A., Štemberger, T., & Klančar, A. (2025). An empirical study on basic and conceptual knowledge, procedural knowledge and problem solving among primary school students. *International Journal of Instruction*, 18(4), 627–650. <https://doi.org/10.29333/iji.2025.18434a>