MORFAI JOURNAL Multidiciplinary Output Research For Actual and International Issue
Multidisciplinary Output Research For Actual and International Issue (MORFAI Journal) ISSN (e): 2808-6635
Volumes 6 No. 6 (2026)

IMPROVING MATHEMATICAL CRITICAL THINKING SKILLS IN FRACTIONS THROUGH THE GASING METHOD BASED ON THE TRADITIONAL MONG-MONGAN BILANGAN GAME: A QUASI-EXPERIMENTAL STUDY OF FIFTH-GRADE ELEMENTARY SCHOOL STUDENTS

Kadek Ashadi Putra[1] , Yeni Yuniarti[2], Yunus Abidin[3]
Department of Elementary Education, Universitas Pendidikan Indonesia, Bandung, Indonesia[1,2,3]
Corresponding author: kadekputra75@guru.sd.belajar.id

Received:05/09/2026 | **Revised** : 12/09/2026 | **Accepted:** 30/09/2026 | **Published:** 19/11/2026

Abstract

Mathematical critical thinking is a key higher-order competence. The fraction training in Indonesian elementary school is generally procedure-oriented, rather than reason-oriented. The present study investigated whether the combination of the traditional game Mong-Mongan Bilangan and the GASING method (concrete-abstract-mental calculation stages) resulted in better gains of mathematical critical thinking than deep learning teaching. A quasi-experimental pretest–posttest nonequivalent control group design was employed with 344 fifth-grade children from six public primary schools in Buleleng Regency, Bali (171 experimental, 173 control), who were taught by trained classroom instructors in seven courses on fractions. Critical thinking was assessed with a 15-item essay test based on Facione's five indicators (interpretation, analysis, assessment, inference, explanation) with scores 1–4 by rubric (Cronbach's alpha = 0.867–0.918). Gain was assessed via normalized gain (N-Gain) and one way MANOVA with univariate follow-up. Pretest scores were similar (50.06 vs 50.26; p=0.907). The mean N-Gain was larger in the experimental group (0.502) than in the control group (0.310), $F(1, 342) = 130.57$, $p < 0.001$, partial $\eta^2 = 0.276$, Cohen's $d = 1.23$. The benefit was observed in all six schools and across all five indicators, however it was not significant at SD Negeri 2 Pemuteran (difference = 0.027; p = 0.557), and the technique × school interaction was significant. The control group likewise improved dramatically . In this case, GASING with Mong-Mongan Bilangan led to greater improvements in mathematical critical thinking than deep learning instruction; however, because intact classrooms were not randomized, the results should not be interpreted as evidence of causation.

Keywords: GASING method; Mong-Mongan Bilangan; mathematical critical thinking; fractions; traditional games; quasi-experiment

1. Introduction

Critical thinking in mathematics is a core higher-order thinking skill in mathematics education. Facione (1990) describes it as "purposeful reflective judgment that results in interpretation, analysis, evaluation, inference, and explanation" (p. This is evident in fifth-grade lessons on fractions, when students explain why a specific step is proper (e.g., why the determinants must be made the same before adding fractions) rather than just following a technique.

Indonesian students' mathematics performance in international exams shows that applying concepts and reasoning is still a challenge. The average TIMSS 2015 score for fourth-grade pupils was 397 (National Center for Education Statistics, n.d.), and the PISA 2025 mathematics score for Indonesian 15-year-olds was 364, compared with the OECD average of 463 (OECD, 2026). Neither assessment evaluates critical thinking directly, so they are utilized here merely as signs that competencies involving application and reasoning need to be reinforced regularly.

This broader picture is reflected in the conditions in the target classroom. The observation of three fraction learning processes in grade V, SD Negeri 2 Pemuteran, demonstrated a pattern in which the teacher presented a procedure, gave one example, and then asked pupils to do exercises independently. Students were seldom prompted to explain the logic of their steps and no activities with specific instructional aids were seen. The first test was carried out on 43 students with an average score of 46.9 out of 100 in mathematical critical thinking. The number of students in the good group was 7 students (16%) and in the bad category was 30 students (70%).

Yohanes Surya's GASING approach (Gampang, Asyik, dan Menyenangkan—Easy, Fun, and Enjoyable) develops ideas via tangible, abstract, and mental-calculation stages (Aksari et al., 2024). Several research have indicated the results of GASING on learning attainment and mathematical thinking skills to be beneficial. Although Rahmawati et al. (2026) utilized a one-group approach, they reported an increase in the mean mathematical thinking score from 59.0 to 82.4 among 37 fourth-grade pupils. Samaudin et al. (2026) reported the improvement of

mathematical critical thinking in 96 fourth-grade students in the Realistic Mathematics Education-based GASING group and the comparison group. Meanwhile, traditional games have been reported to be contributory to numeracy literacy and critical thinking (Ni'mah et al., 2024; Handarizki et al., 2025). The activities, learning material and designs were different throughout those research, thus their findings are used as a supporting backdrop, rather than particular evidence for Mong-Mongan Bilangan.

Mong-Mongan is a guessing game that is popular in Bali. This study modified it to Mong-Mongan Bilangan for fraction learning and played in the mental-calculation stage of GASING. Students worked in small groups and rolled fraction dice. One participant acted as banker of the game, rolling the dice and judging responses. The players that answered correctly got star points, and the group with the highest star points was the winner. In theory, the duty of the banker to evaluate answers and the need to answer in front of the group under a time limit provides opportunity to practice the evaluation of answers, forming conclusions and explaining reasoning.

The literature search that was undertaken for this study did not find a verifiable GASING study that evaluated GASING integrated with a traditional game versus a comparison instructional technique for fraction learning; this research gap assessment is based on a limited search. Thus, this paper intends to investigate the difference of improving mathematical critical thinking skills between fifth-grade students who learn with the GASING Method using the traditional Mong-Mongan Bilangan game and students who learn with deep learning teaching. The hypothesis examined was that the mathematics critical thinking skills improvement would be different between the two groups with a higher improvement in the experimental group.

2. Literature Review

2.1 The GASING Method and Its Stages

GASING is a learning method in mathematics which makes mathematics easy, engaging, and pleasurable to grasp in methodical, incremental phases (Aksari et al., 2024). The procedure consists of three consecutive steps. In the concrete stage, students work with real items or visual representations to make abstract concepts more tangible (Ebner et al., 2025). In the abstract stage, pupils move to mathematical symbols and notation. In the mental-calculation stage, pupils calculate rapidly and accurately without the use of aids, showing that they have internalized their understanding.

This pattern is congruent with constructivist theory, which views knowledge as the consequence of learners' active construction (Piaget, 1964; Vygotsky, 1978), and with Bruner's (1966) three forms of representation: enactive, iconic, and symbolic. The "easy" principle, which breaks down content into little steps, is congruent with scaffolding within the zone of proximal development (Vygotsky, 1978). Skemp (1978) distinguishes between instrumental and relational awareness; advancing too rapidly to the abstract stage may result in procedural memorization without conceptual understanding.

2.2 The Traditional Mong-Mongan Bilangan Game

Traditional games are cultural goods which emerge in certain groups and bear cultural values handed on from generation to generation. Ethnomathematics research have identified mathematical ideas in such games, such as addition, subtraction, and mixed fractions during the Bang-Jo-Neng-Teh-Ru game (Rosadi, 2024). Mong-Mongan is a guessing game from Bali that involves social interaction and easy rules. In this study, the game was modified to Mong-Mongan Bilangan, employing fraction dice, a banker, star points, two to three rounds, and a five-minute time constraint and was played during the mental-calculation stage.

2.3 Mathematical Critical Thinking

Critical thinking is an active, reflective process of judgment that involves interpretation, analysis, assessment, inference, and explanation, as well as self-regulation (Facione, 1990). The first five signals were included in this study because they can be quantified by utilizing essay questions that are scored using a rubric, whereas self-regulation is not easily measured directly through a written test for primary school students. Interpretation is the comprehension of the significance of the image or circumstance in the form of a fraction. Analysis is the analysis of relationships and criteria such as the requirement that the two portions must be equal in size. Evaluation is to determine the accuracy of a solution and remedy the error . Inference is the process of drawing conclusions from the results of calculations and explanation is the process of expressing the rationale for each step in a logical sequence.

OPEN ACCESS

2.4 Previous Studies

The GASING research evaluated typically reported beneficial findings. Nurjanna et al. (2016) showed an average gain of 0.42 on two-digit multiplication using a one-group strategy. Buknoni et al. (2026) found that the N-Gain of GASING was 0.83, while the N-Gain of direct instruction was 0.63 among junior high school students. Yusuf et al. (2026) showed an N-Gain of 0.77 for dice-game-based GASING versus 0.35 for traditional education in elementary students' multiplication skills. Pratiwi et al. (2026) conducted a comprehensive review, which showed that GASING generally enhances conceptual understanding and student participation, however most research are short-term. Concerning the traditional game, Ni'mah et al. (2024) found that Balogo game-based learning led to better improvements in numeracy literacy and critical thinking than conventional learning.

These studies vary in construct, topic, educational level, and comparative condition, therefore they are utilized as background rather than as proof specifically for Mong-Mongan Bilangan. At the time of the literature search for this study, no study was located that paired GASING with the classic Mong-Mongan game for fraction learning, and that included a comparison group.

3. Method

3.1 Research Design

This study used a quantitative technique with a quasi-experimental pretest-posttest non-equivalent control group design. This strategy was used because the participants were intact classes and individual randomization was impossible (Fraenkel et al., 2019). The experimental group was taught by using the GASING Method with the traditional Mong-Mongan Bilangan game while the control group was taught by deep learning instruction, namely student-centered instruction emphasizing meaningful understanding without the GASING stages and the Mong-Mongan Bilangan game. Both groups studied the same topic and received the same teaching time allocation.

3.2 Participants and Group Assignment

The study population consisted of 344 fifth-grade children from six A-accredited public primary schools in Buleleng Regency, Bali, during the second semester of the 2025/2026 academic year. The sample consisted of 171 students in the experimental group and 173 in the control group (Table 1). Two parallel fifth grade classes from each school were matched on average mathematics scores from the previous semester, class size and school characteristics. The mean scores of experimental and control groups are 80.18 and 80.17 in mathematics correspondingly, a difference of just 0.01 points. In all six schools the group assignment was a fixed rule; Class V A was the control group and Class V B was the experimental group. This was not a random assignment, therefore equivalency was not assumed and was checked with pretest scores.

Table 1. Distribution of Research Participants

School	Experimental	Control	Total
SD Negeri 2 Pemuteran	21	22	43
SD Negeri 3 Seririt	34	34	68
SD Negeri 1 Paket Agung	30	31	61
SD Negeri 1 Banjar Jawa	41	42	83
SD Negeri 3 Sukasada	24	23	47
SD Negeri 1 Banyuning	21	21	42
Total	**171**	**173**	**344**

3.3 Intervention and Procedure

The instruction was conducted by fifth-grade classroom instructors at each school, not the researchers. Teachers in the experimental classes were trained in GASING Method for fraction content, and teachers in the control classes were trained in teaching deep learning. In March–April 2026, each class participated in nine sessions: one pretest session (120 minutes), seven teaching sessions (2 × 35 minutes) and one posttest session (120 minutes). The seven lessons included learning about fractions, equivalent fractions, simplifying fractions, and adding and subtracting fractions with like and unlike denominators. at the experimental group each session was conducted with the sequence of concrete, abstract, and mental-calculation, worksheet discussion and reflection, and the Mong-Mongan Bilangan game at the mental-calculation stage (two to three rounds, five minutes).

3.4 Instrument

Mathematical critical thinking skills were tested using an essay test created from five indications of Facione (1990). The questions did not contain GASING-specific media or games to maintain fairness to both groups. Eighteen questions were pilot tested and 15 were selected for the final test with three questions for each indicator. Each question was graded from 1 to 4 based on a rubric. The overall score ranged from 15 to 60 and was translated to a 0–100 scale by applying the formula (obtained score ÷ 60)*100. The scoring categories were very good (86-100), good (71-85), reasonable (56-70), and low (< 56).

Content validity was evaluated by two experts (a university lecturer and a mathematics teacher) using Gregory's coefficient (2011), and all 18 questions obtained V = 1.00. The pilot test was conducted on February 23, 2026, with 30 students at SD Negeri 7 Kampung Baru who were not part of the sample. The results revealed that 17 of the 18 questions were legitimate (calculated r = 0.408-0.772; r-table = 0.361). Question 17 was invalid (r = 0.104). To reach the planned number of questions, the two questions with the lowest corrected r values (Questions 18 and 1) were dropped as valid questions. The final exam thus consisted of pilot-test Questions 2–16. The Cronbach's Alpha coefficient for the final questions in the pilot test was 0.887. In the study data (n = 344) it was 0.918 for the pretest and 0.867 for the posttest.

3.5 Data Analysis

Improvement was measured using normalized gain (Hake, 1998): g = (posttest score − pretest score) ÷ (100 − pretest score), with low (g < 0.30), moderate (0.30 ≤ g ≤ 0.70), and high (g > 0.70) categories. Because the groups were not perfectly matched, N-Gain was examined (Bao, 2006; Hake, 1998). Assumptions were evaluated with Shapiro-Wilk (normality), Box's M (equality of covariance matrices) and Levene's test (equality of variances). The hypothesis was assessed by one-way MANOVA on N-Gain scores for critical thinking and self-confidence, followed by univariate testing (testing of Between-Subjects Effects) at the 0.05 level of significance. The article gives the findings for critical thinking. Effect sizes were given using partial eta-squared and Cohen's d (Cohen, 1988). Additional analyzes consisted of an independent-samples t-test on pretest scores, paired-samples t-tests, the Mann–Whitney U test with a Bonferroni correction (α = 0.025), and MANOVA with instructional technique, school, and their interaction as factors. All analyzes were performed using SPSS.

4. Results

4.1 Baseline Equivalence and Descriptive Statistics

The two groups were matched on baseline abilities. Experimental group mean pretest score = 50.06, control group mean pretest score = 50.26 (difference = −0.20; 95% CI [−3.60, 3.19]; t(342) = −0.117; p = 0.907). Mean posttest scores of the experimental group and control group were 74.39 and 65.78 respectively after the intervention. The mean gain was 24.33 points in the experimental group and 15.52 points in the control group (Table 2).

Table 2. Mathematical Critical Thinking Scores and N-Gain by Group

Group	n	Pretest (SD)	Posttest (SD)	N-Gain (SD)	N-Gain min–max
Experimental	171	50.06 (16.07)	74.39 (12.34)	0.502 (0.147)	0.167–0.929
Control	173	50.26 (15.94)	65.78 (13.73)	0.310 (0.164)	−0.071–0.741

Note: SD = standard deviation; scores are on a 0–100 scale.

In the pretest, majority of the students were in the poor category with 110 students (64.3%) in the experimental group and 111 students (64.2%) in the control group. The posttest results revealed that the experimental group was dominated by the good (70 students; 40.9%) and very good (37 students; 21.6%) categories, while only 13 students (7.6%) remained in the low category. In the control group, 43 kids (24.9%) were still in the low category and just 12 students (6.9%) were in the very good category.

4.2 Improvement (N-Gain)

The average N-Gain of the experimental group (0.502) was in the medium category and higher than the control group (0.310) which was at the lower boundary of the medium category. The control group comprised two children with negative N-Gain values, and data from these kids were kept as legitimate measurement results. In the experimental group, 158 kids (92.4%) reached medium or high N-Gain, while 91 students (52.6%) in the control group were still in the low category (Table 3).

Table 3. Distribution of Mathematical Critical Thinking N-Gain Categories

N-Gain Category	Experimental	Control
High ($g > 0.70$)	16 (9,4%)	1 (0,6%)
Medium ($0.30 \leq g \leq 0.70$)	142 (83,0%)	81 (46,8%)
Low ($g < 0.30$)	13 (7,6%)	91 (52,6%)
Total	171 (100%)	173 (100%)

The improvement was larger in the experimental group than in the control group for all five indicators and was relatively uniform across indicators (Table 4). Indicator N-Gain values of the experimental group varied from 0.477 to 0.509, and the control group had values ranging from 0.306 to 0.326. Differences between indicators were modest and were not tested inferentially, and so should be read descriptively.

Table 4. Achievement by Critical Thinking Indicator (0–100 scale; N-Gain calculated from group means)

Indicator	Experimental pretest	Experimental posttest	Experimental N-Gain	Control pretest	Control posttest	Control N-Gain
Interpretation	49,3	73,4	0,477	50,4	66,0	0,313
Analysis	49,4	74,1	0,489	50,5	65,8	0,308
Evaluation	51,5	76,1	0,509	50,5	65,7	0,306
Inference	50,3	74,3	0,483	49,8	66,2	0,326
Explanation	49,9	73,8	0,479	50,2	65,5	0,307

4.3 Assumption Testing

The Shapiro-Wilk test showed that the critical thinking N-Gain was normally distributed in the experimental group ($W=0.993$; $p=0.532$) and control group ($W=0.986$; $p=0.071$). Box's M test suggested that covariance matrices were homogenous ($M = 2.947$; $p = 0.402$) and Levene's test indicated that variances were homogeneous ($F(1, 342) = 3.507$; $p = 0.062$). The p-values of Levene's test and normality test of the control group were near to the threshold of 0.05 . However, the sizes of the groups were about equal (171 and 173) and MANOVA is resistant to small variations . The Kolmogorov–Smirnov test with Lilliefors correction for the control group was $p \approx 0.047$. This, however, did not change the choice as skewness and kurtosis were within permissible limits and nonparametric test resulted in the same conclusion.

4.4 Hypothesis Testing

The MANOVA revealed a significant simultaneous difference of critical thinking and self-confidence N-Gain (Wilks' Lambda = 0.677, $F (2, 341) = 81.114$, $p < 0.001$, partial η 2 = 0.323). The univariate test revealed that instructional approach significantly affected critical thinking N-Gain, $F(1, 342) = 130.572$; $p < 0.001$; partial $\eta 2 = 0.276$. The effect size (Cohen's d) was 1.23, indicating a strong effect (Cohen, 1988), with the direction of the difference favoring the experimental group. So the null hypothesis was rejected and the research hypothesis was accepted.

Further analyzes led to the same conclusion. Paired-samples t-tests showed that posttest scores were considerably higher than pretest scores in the experimental group (difference=24.33; $t(170)=34.729$; $p<0.001$; $dz=2.66$) as well as in the control group (difference=15.52; $t(172)=21.237$; $p<0.001$; $dz=1.61$). The improvement in the experimental group so was above the improvement in the control group. The Mann-Whitney U test was statistically significant at the Bonferroni adjusted threshold of 0.025 (Mean ranks = 225.12 experimental group vs. 120.48 control group; $U = 5793.5$; $Z = -9.757$; $p < 0.001$).

4.5 Consistency Across Schools

The mean N-Gain of the experimental group was better than that of the control group in all six schools (Table 5). Significant differences were identified in five schools at the 5% level, however at SD Negeri 2 Pemuteran the difference was just 0.027 and was not significant ($p = 0.557$). The school-level tests show the consistency in the direction of the effect, but are not the major hypothesis tests because the sample sizes at each school were tiny.

OPEN ACCESS

Table 5. Mean Mathematical Critical Thinking N-Gain at Each School

School	Experimental	Control	Difference	p (Welch's t-test)
SDN 2 Pemuteran	0,507	0,480	0,027	0,557
SDN 3 Seririt	0,576	0,287	0,288	< 0,001
SDN 1 Paket Agung	0,368	0,203	0,165	< 0,001
SDN 1 Banjar Jawa	0,476	0,266	0,211	< 0,001
SDN 3 Sukasada	0,600	0,470	0,130	< 0,001
SDN 1 Banyuning	0,509	0,239	0,270	< 0,001

A subsequent MANOVA including instructional technique, school and their interaction as factors indicated that the effect of instructional method remained significant after controlling for school (Wilks' Lambda = 0.587; F(2, 331) = 116.548; $p < 0.001$; partial η 2 = 0.413). There was a significant technique × school interaction (Wilks' Lambda = 0.883; F(10, 662) = 4.238; $p < 0.001$; partial $\eta^2 = 0.060$), suggesting that the size of the experimental group advantage varied among schools. No school reversed course on the advantage.

5. Discussion

The learning process was not examined in this study, but the outcomes were measured before and after the intervention. Thus, explanations for the disparities are theoretical interpretations rather than findings of examinations of the underlying systems.

5.1 Main Findings

The mathematics critical thinking N-Gain of the experimental group was 0.502 or 0.192 greater than the control group (0.310) with a large effect size (Cohen's d = 1.23; partial η 2 = 0.276). The two groups were equal at pretest (50.06 and 50.26), hence the difference in posttest scores (74.39 and 65.78) was not a reflection of measured differences in ability at baseline. The control group also improved significantly (15.52 points; dz = 1.61) , demonstrating that deep learning instruction itself was followed with an increase in critical thinking, and that the experimental group's advantage represented an additional gain beyond that improvement.

5.2 Theoretical Explanation

One reason that is consistent with the literature review is the sequence of stages in GASING. The test questions required students to comprehend diagrams of fractions, analyze circumstances and relationships between numerator and denominator, evaluate flaws in solutions, draw conclusions, and explain their reasoning. Students who had previously interacted with fraction objects and diagrams in the concrete stage had representations to fall back on to meet these needs, whereas training that did not include that step might have depended more on processes. This is in line with the instrumental and relational knowledge (Skemp, 1978) and enactive-iconic-symbolic sequence (Bruner, 1966) dichotomy.

The game part, too, might have had something to do with it. In Mong-Mongan Bilangan, the banker judges the other players' answers and the players must answer in front of the group in a given time limit. This in theory is an opportunity to practice judging the correctness of replies and communicating logic, directly tied to the assessment and explanation indications. This theory is also supported by the descriptive data but does not prove it: evaluation in the experimental group had the highest N-Gain (0.509), but all indicators were in a narrow range of 0.477–0.509. This equal pattern of improvement more supports the conclusion that the intervention increased critical thinking in general rather than any one metric.

5.3 Comparison with Previous Studies

The results are in line with Samaudin et al. (2026) who showed an improvement of mathematics critical thinking of fourth-grade children who received GASING training, and Rahmawati et al. (2026) who observed an increase in mathematical thinking in one-group design. One-group studies cannot separate the intervention impact from other factors, and hence the present study, with a control group, gives greater evidence. The difference in the present study (0.19) was similar to the difference reported by Buknoni et al. (2026; 0.83 and 0.63, a difference of 0.20) and smaller than the difference reported by Yusuf et al. (2026; 0.77 and 0.35, a difference of 0.42), whereas the absolute N-Gain of the experimental group (0.50) was lower than in both studies. Caution should be exercised in making these

OPEN ACCESS

comparisons as there were differences in measured constructs (essay-based critical thinking versus calculation skills or conceptual comprehension), educational levels and content. Furthermore, the comparator condition in this study was deep learning, which may be stronger than direct or traditional instruction.

Ni'mah et al. (2024) and Handarizki et al. (2025) observed a larger improvement in critical thinking through traditional-game-based learning, which supports the traditional-game component. The games, content, and research procedures in those studies were different, so their findings are presented as a supporting backdrop, not as evidence for Mong-Mongan Bilangan. Neither did this study compare GASING without the game to GASING with the game, thus the contributions of the two components cannot be disentangled.

The mean of N-Gain of experimental group was in medium category, not in high category, and only 16 students (9.4%) got high category score. Critical thinking is a higher-level skill that takes time to develop, whereas the intervention was just seven lessons. Pratiwi et al. (2026) stated that most studies on GASING are short-term, and this restriction applies here.

5.4 Variation Across Schools

The direction of the experimental group advantage was comparable throughout the six schools, but its magnitude varied: the N-Gain difference ranged from 0.027 at SD Negeri 2 Pemuteran to 0.288 at SD Negeri 3 Seririt, and the technique × school interaction was significant. The control group of SD Negeri 2 Pemuteran got an N-Gain of 0.480, far higher than the control-group results in the other schools (0.203–0.470) which resulted in no significant difference. No data were collected in this investigation that could explain the cause. The possible explanations are the very active use of deep learning in that class, the differences in the quality of teachers' implementation, or the small class numbers (21 and 22 pupils). Low baseline ability alone cannot be the sole explanation for the outcome, because SD Negeri 3 Sukasada also had a low mean pretest score and nevertheless exhibited a difference of 0.130. This result is a warning against generalizing the results as an effect of identical magnitude in every class.

5.5 Implications and Limitations

Practically, these findings can be used as a basis for fifth-grade instructors to try the concrete-abstract-mental calculation sequence with the Mong-Mongan Bilangan game in fraction lessons, including a student worksheet and reflection in each session. As the effect magnitude differed between schools, teachers are encouraged to assess learning results in their own courses. The limitations of this study are the following.

1. The quasi-experimental design did not allow for random assignment. By a predetermined rule, Class V A was assigned as the control group and Class V B as the experimental group; thus, class order effect cannot be ruled out. Each school supplied just one class per condition, therefore intervention effects cannot be disentangled from teacher and class effects. Thus, the term "effect" should be understood to mean a difference in improvement compatible with the intervention, rather than definite evidence of causality.
2. Students were clustered within six schools, however the main study regarded students as independent units. The issue was decreased by the extra study including school as a factor, however 6 schools were too few for an adequate multilevel model.
3. There was no data regarding intervention fidelity, for example if deep learning training in the control classes was implemented as designed.
4. The posttest was conducted after seven lessons and there was no follow-up measurement, such that the permanence of the improvement is unknown.
5. The content was restricted to fifth grade fractions, therefore no generalization to other topics or educational levels can be made yet.
6. The contributions of GASING and the Mong-Mongan Bilangan game cannot be detached since there was no group that experienced only one of the two aspects.

6. Conclusion and Recommendations

The GASING Method using the traditional Mong-Mongan Bilangan game among 344 fifth-grade students at six elementary schools in Buleleng Regency was more effective in improving mathematical critical thinking in fraction content than deep learning instruction (N-Gain = 0.502 vs 0.310; $F(1, 342) = 130.57$; $p < 0.001$; partial $\eta^2 = 0.276$; Cohen's $d = 1.23$). The advantage was present across all five of Facione's indicators and at all six schools, but was not significant at SD Negeri 2 Pemuteran, and its size varies across schools. The control group also improved, thus this advantage is relative to teaching in deep learning. Because intact classrooms were not randomized, this

conclusion suggests a difference in outcomes between two instructional styles under the settings of this study rather than definite evidence of causation.

Recommendations for instructors and school. Teachers can use the concrete–abstract–mental calculation process with the Mong-Mongan Bilangan game in fraction lessons and accompany this with reasoning questions that develop analysis, assessment, and drawing conclusions. Schools should provide teacher training, gaming materials, and oversee the application of teaching to maintain consistency throughout the courses.

Recommendations for researchers in the future. Future studies should (1) randomly assign or control for covariates of baseline ability and class characteristics; (2) monitor intervention fidelity; (3) lengthen the intervention and include follow-up measurements; (4) disentangle the contributions of GASING and the Mong-Mongan Bilangan game by using separate comparison groups; and (5) apply the approach to other content areas and educational levels.

REFERENCES

Aksari, I. H., Fathurrohman, I., & Fajrie, N. (2024). Implementasi metode pembelajaran GASING untuk meningkatkan hasil belajar siswa. *Jurnal Prakarsa Paedagogia, 7*(2), 32–42. https://doi.org/10.24176/jpp.v7i2.12978

Bao, L. (2006). Theoretical comparisons of average normalized gain calculations. *American Journal of Physics, 74*(10), 917–922. https://doi.org/10.1119/1.2213632

Bruner, J. S. (1966). *Toward a theory of instruction*. Belknap Press of Harvard University Press.

Buknoni, Y., Wahyudin, W., & Iriawan, S. B. (2026). Efektivitas metode GASING untuk meningkatkan pemahaman konsep bilangan bulat di sekolah. *Scholaria: Jurnal Pendidikan dan Kebudayaan, 16*(2), 194–203. https://doi.org/10.24246/j.js.2026.v16.i2.p194-203

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Ebner, S., MacDonald, M. K., Grekov, P., & Aspiranti, K. B. (2025). A meta-analytic review of the concrete-representational-abstract math approach. *Learning Disabilities Research and Practice, 40*(1), 31–42. https://doi.org/10.1177/09388982241292299

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction* (Research findings and recommendations). American Philosophical Association. https://eric.ed.gov/?id=ED315423

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2019). *How to design and evaluate research in education* (10th ed.). McGraw-Hill Education.

Gregory, R. J. (2011). *Psychological testing: History, principles, and applications* (6th ed.). Pearson.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*(1), 64–74. https://doi.org/10.1119/1.18809

Handarizki, N. L., Setiawan, A., & Yunitasari, I. (2025). The effect of traditional engklek game based on problem based learning model on improving critical thinking ability of elementary school students in science subjects. *JOSAR (Journal of Students Academic Research), 10*(2), 177–188. https://doi.org/10.35457/pmsxqx47

National Center for Education Statistics. (n.d.). *TIMSS - Mathematics for grades 4 and 8: Averages* (Table 1. Average mathematics scores of 4th-grade students, by education system: 2015). U.S. Department of Education. https://nces.ed.gov/timss/timss2015/timss2015_table01.asp

Ni'mah, N., Setyawan, D., & Astuti, A. D. (2024). Efektivitas pembelajaran berbasis permainan tradisional "Balogo" terhadap literasi numerasi dan berpikir kritis pada siswa sekolah dasar. *Pedagogik: Jurnal Pendidikan, 19*(2), 207–214. https://doi.org/10.33084/pedagogik.v19i2.8400

Nurjanna, S., Januarsih, J., Pratama, R. A., & Sulistiawati, S. (2016). Pengaruh pembelajaran matematika GASING (Gampang, ASyIk, dan menyenangkaN) pada materi perkalian bilangan dua angka dengan bilangan dua angka terhadap hasil belajar siswa kelas III A Sekolah Dasar Negeri Cihuni 1, Kelapa Dua, Tangerang. In *PRISMA, Prosiding Seminar Nasional Matematika IX 2015* (pp. 614–621). Jurusan Matematika FMIPA Universitas Negeri Semarang. https://journal.unnes.ac.id/sju/prisma/article/view/21701

OECD. (2026). *PISA 2025 results (Volume I): Indonesia* [Country note]. OECD Publishing. https://www.oecd.org/en/publications/pisa-2025-results-volume-i-country-notes_2d4ff9ea-en/indonesia_9c880f8a-en.html

Piaget, J. (1964). Part I: Cognitive development in children: Piaget development and learning. *Journal of Research in Science Teaching, 2*(3), 176–186. https://doi.org/10.1002/tea.3660020306

Pratiwi, Y., Herlambang, Y. T., Yuniarti, Y., & Bakar, A. (2026). Transformation of elementary school students' mathematics learning outcomes using the GASING method in a critical pedagogical approach. *Journal of Educational Sciences, 10*(4), 195–206. https://doi.org/10.31258/jes.10.4.p.195-206

OPEN ACCESS

Rahmawati, R. S. A., Sopandi, W., & Turmudi, T. (2026). Metode GASING (Gampang, Asyik, dan Menyenangkan) dalam meningkatkan kemampuan berpikir matematis siswa sekolah dasar. *Jurnal Yudistira: Publikasi Riset Ilmu Pendidikan dan Bahasa, 4*(4), 1–9. https://doi.org/10.61132/yudistira.v4i4.3020

Rosadi, M. R. I. (2024). Eksplorasi konsep penjumlahan, pengurangan dan pecahan campuran pada permainan Bang-Jo-Neng-Teh-Ru. *Maliki Interdisciplinary Journal, 2*(4), 417–424. https://urj.uin-malang.ac.id/index.php/mij/article/view/4452

Samaudin, L., Irianto, D. M., & Jenuri, J. (2026). The effectiveness of the PMR-based GASING method in enhancing elementary school students' critical thinking skills. *Jurnal Pembelajaran, Bimbingan, dan Pengelolaan Pendidikan, 6*(10), 13. https://doi.org/10.17977/um065.v6.i10.2026.13

Skemp, R. R. (1978). Relational understanding and instrumental understanding. *Arithmetic Teacher, 26*(3), 9–15.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.

Yusuf, D. A., Muslihin, H. Y., & Apriani, I. F. (2026). Effectiveness of the GASING method based on dice games in improving multiplication skills of primary school students. *Indonesian Journal of Primary Education, 10*(1), 167–186. https://doi.org/10.17509/ijpe.v10i1.98405

OPEN ACCESS