Jurnal Pendidikan Progresif
e-ISSN: 2550-1313 | p-ISSN: 2087-9849
URL: https://jpp.fkip.unila.ac.id/index.php/jpp
Jounal's DOI: http://doi.org/10.23960/jpp

Vol. 16, No. 4, 2026
Article ID: 1990
Pages: 2155-2168

When Learning Becomes a Game: The Impact of Ladders War in Game-Based Learning on Students' Mathematical Critical Thinking Skills

Kamiliya Nailah Fitri[1], Hafiziani Eka Putri[1,*], Alfiana Nurussama[1], Mariano Dos Santos[2], & Sri Hastuti Noer[3]

[1]Elementary School Teacher Education Study Program, Universitas Pendidikan Indonesia, Indonesia
[2]Instituto Católico Para A Formação De Professores (ICFP) Baucau, Timor-Leste
[3]Mathematics Education Study Program, Universitas Lampung, Indonesia

*Corresponding email: hafizianiekaputri@upi.edu

Received: 03 September 2026 *Accepted: 20 October 2026* *Published: 04 November 2026*

Abstract: Mathematical critical thinking remains a persistent challenge in elementary education globally, as TIMSS 2023 findings show that only a small proportion of fourth-grade students worldwide reached the High and Advanced benchmarks in mathematics, while empirical evidence on the effectiveness of game-based approaches at the primary level remains limited. This study examined how the Game-Based Learning (GBL) model, assisted by Ladders War media, influenced fourth-grade students' mathematical critical-thinking skills in number patterns, a topic fundamental to algebraic reasoning that remains persistently difficult for young learners. Employing a quasi-experimental Nonequivalent Control Group Design, 50 students (aged 9–10; 25 males, 25 females) from two intact classes in Bekasi, Indonesia, participated. The study was conducted over four 90-minute sessions: the pretest was administered at the beginning of the first session, followed by the first treatment, two additional treatment sessions, and a posttest in the fourth session. Critical thinking was assessed using a five-item expert-validated essay test (reliability = 0.74) measuring analysis, inference, evaluation, and explanation. ANCOVA with pretest scores as the covariate yielded a large effect size (ç²p = 0.217) and a significant between-group difference ($F = 13.013$, $p < .001$), with an adjusted mean difference of 13.132 points (95% CI [5.809, 20.455]). The GBL group achieved a higher learning gain (N-Gain = 0.56, moderate) than the control group (N-Gain = 0.33, moderate). These findings suggest that the GBL model with Ladders War media effectively supports the development of mathematical critical thinking among elementary school students. Theoretically, this study reinforces the use of GBL to foster higher-order thinking, while practically offering an innovative instructional alternative for mathematics teachers, with broader implications for mathematics education in developing-country contexts.

Keywords: game-based learning, mathematical critical thinking, ladders war.

Article's DOI: https://doi.org/10.23960/jpp.v16i4.pp2155-2168

INTRODUCTION

Mathematical education plays an important role in developing students' problem-solving, analytical, and logical reasoning skills, which are essential components of 21st-century competencies (Thornhill-Miller et al., 2023). Mathematics provides a foundation across scientific disciplines and fosters students' logical and critical thinking by developing structured, abstract, and systematic concepts grounded in deductive reasoning (Sukmawati & Arifanti, 2022). This aligns with the National Council of Teachers of Mathematics (2000), which asserts that effective mathematics instruction must foster conceptual connections, the communication of mathematical ideas, and robust reasoning. Nevertheless, students' mathematical critical thinking skills remain a significant challenge that classroom practice has yet to address fully.

Mathematical critical thinking comprises a range of cognitive abilities that allow students to analyze information systematically, evaluate the

validity of mathematical arguments, derive logical conclusions, and construct well-reasoned justifications when addressing mathematical problems (Facione, 2011). In mathematics learning, this ability not only supports conceptual mastery but also encourages students to identify relationships among concepts, evaluate problem-solving strategies, construct evidence-based conclusions, and communicate mathematical reasoning logically (Jannah et al., 2024; Monteleone et al., 2023). These abilities align with the characteristics of mathematics, which emphasize logical reasoning, systematic analysis, and problem-solving as the foundation for building conceptual understanding. Therefore, students should develop mathematical critical-thinking skills at the elementary level to build a solid foundation for addressing increasingly complex problems in later stages of education.

Cultivating mathematical critical-thinking skills among elementary school students remains a persistent, well-documented challenge in mathematics education worldwide. TIMSS 2023 data show that only a limited proportion of fourth-grade students achieved the High and Advanced International Benchmarks in mathematics, reflecting limited cultivation of higher-order mathematical reasoning competencies at the primary school level (IEA, 2023). This aligns with Monteleone et al. (2023), who found that many elementary school students still struggle to analyze, evaluate, and communicate mathematical reasoning systematically. One learning context that exemplifies this challenge is number pattern material, which this study selected because it forms the foundation of algebraic thinking. This topic requires students to recognize regularities, identify relationships between patterns, generalize, and construct general rules based on observed patterns, making it well suited to developing mathematical critical thinking skills at the elementary level (Maula et al., 2024; Pramasdyahsari et al., 2023). A similar challenge exists in Indonesia, where mathematics instruction remains dominated by procedurally oriented approaches, leaving the cultivation of higher-order thinking skills suboptimal.

Despite the growing adoption of Game-Based Learning in elementary mathematics education, recent reviews indicate that research gaps remain regarding mathematical topics, research designs, game types, and empirical evidence across diverse learning contexts, highlighting the need for further research into its effectiveness on students' higher-order thinking skills (Alotaibi, 2024; Dan et al., 2024; Wardani & Himmah, 2026). These gaps underscore the need for research that develops innovative game-based media within the Game-Based Learning framework, specifically designed to develop mathematical critical thinking skills in number pattern material at the elementary school level.

Game-Based Learning (GBL) is a learning framework that integrates game elements into academic tasks to create meaningful learning experiences and support students' cognitive development in formal educational settings (Plass et al., 2015). This study's implementation of GBL is grounded in Vygotsky's (1978) constructivist perspective, which views knowledge as constructed through active experience and social interaction. These principles are reflected in problem-solving, discussion, decision-making, and peer interaction during gameplay. These activities give students opportunities to develop mathematical understanding through structured learning experiences.

Games in mathematics education also rest on a strong theoretical foundation. Dienes (1963) emphasized the importance of active, game-based experiences in helping students explore and generalize mathematical structures. Consistent with this perspective, well-designed mathematical games can provide an environment for exploration, reasoning, discussion, and the development of conceptual understanding

(Klooger, 2026). In this study, Ladders War serves as a structured game-based medium that allows students to repeatedly solve number-pattern problems, discuss strategies, and make decisions based on their solutions. These activities relate to mathematical critical thinking skills, particularly analysis, inference, evaluation, and explanation, which the test assesses. Accordingly, this study's theoretical foundation focuses on knowledge construction and cognitive development relevant to mathematical critical thinking rather than on student motivation or engagement, as these constructs were not directly measured.

International empirical studies have demonstrated that Game-Based Learning positively influences higher-order thinking skills, mathematical conceptual understanding, learning motivation, and student engagement. These findings are further supported by meta-analyses and systematic literature reviews reporting the effectiveness of Game-Based Learning across various educational levels and mathematics learning contexts (Anggoro et al., 2024; Cheng et al., 2023; Christopoulos et al., 2024; Mao et al., 2022; Putra et al., 2024; Rohman & Khaliza, 2025). However, most existing studies have focused on affective or motivational outcomes, with relatively little attention given to the direct impact of Game-Based Learning on mathematical critical thinking, particularly at the elementary level and within specific content areas such as number patterns. The present study seeks to address this gap.

Various educational game media have been developed in mathematics learning, such as modified Snakes and Ladders games designed to support the learning process and improve student outcomes (Azizah & Fidrayani, 2024; Halidah et al., 2024). However, game media designed specifically to stimulate mathematical critical thinking through decision-making activities and problem-solving strategies remain relatively limited. To address this gap, this study developed Ladders War as a mathematics learning medium by adapting selected game mechanics inspired by the Ladders War challenge featured in the second season of Clash of Champions. The study adapted and redesigned these mechanics to suit elementary school students and the learning objectives of number-pattern material.

These game mechanics were later adapted into a mathematics learning medium suited to elementary school students and the learning objectives of number-pattern material, integrating critical-thinking tasks, action cards, inter-group battle mechanisms, and strategic decision-making to stimulate students' cognitive engagement. Unlike most educational games based on modified Snakes and Ladders, which retain conventional mechanics in which dice rolls determine piece movement (Azizah & Findrayani, 2024), Ladders War uses step cards with number-pattern tasks. Students must solve the tasks correctly to obtain the opportunity to move forward. The ladder mechanism is also modified into an inter-group battle that requires students to solve mathematical problems before advancing. Through these mechanics, the game directly integrates problem-solving and decision-making, positioning Ladders War not merely as a game medium but as a structured learning tool designed to support the development of mathematical critical-thinking skills.

This study's novelty lies in integrating the Ladders War game media into the Game-Based Learning model and assessing its effectiveness across four mathematical critical-thinking indicators (analysis, inference, evaluation, and explanation) within number pattern material at the elementary level, a combination that prior research has rarely examined. Grounded in Constructivism and responding to persistent global concerns about students' higher-order thinking in mathematics, this study aims to contribute not only to the development of

elementary mathematics instruction in Indonesia but also to enrich the international literature on contextual Game-Based Learning implementation for developing mathematical critical thinking across diverse educational systems.

Accordingly, this study hypothesizes that students receiving Game-Based Learning instruction assisted by Ladders War media will demonstrate significantly higher mathematical critical thinking skills than those receiving Discovery Learning with Question Pull media, after statistically controlling for prior ability through ANCOVA (H). The findings are expected to contribute to theory and practice in game-based mathematics instruction at the elementary level, while giving teachers a reference for designing learning experiences that more effectively develop students' mathematical critical-thinking skills.

METHOD

This study employed a quantitative quasi-experimental design with a Nonequivalent Control Group arrangement, in which participants were allocated to experimental and control groups without random assignment (Widiastika et al., 2023). This design was selected because the participating school's intact class structures precluded random assignment of students to different conditions, a common constraint in naturalistic educational research (Ballance, 2024). Although the lack of randomization limits causal inference compared with a true experiment, ANCOVA with pretest scores as a covariate was used to control for pre-existing differences in prior ability between groups, producing a more accurate estimate of the treatment effect when analytical assumptions are met (Lüdtke & Robitzsch, 2025). Prior to data collection, the study received an official research permit from the university (No. T-2026/UN40.C4.1/PK.03.03/2026) and formal approval from the school principal (No. 421.2/044/SDN TP VIII/V/2026). Parents and guardians of all participating students were informed of the research objectives, procedures, and timeline through the classroom teacher via the established parent communication group. All student responses were analyzed anonymously to ensure participant confidentiality.

The participants were 50 fourth-grade students, approximately 9-10 years of age, from two intact classes at a state elementary school in Bekasi City, West Java, Indonesia. The experimental group included 25 students (13 males, 12 females), and the control group included 25 students (12 males, 13 females). Participants were selected through purposive sampling based on class availability and teacher willingness (Ballance, 2024). The experimental group received instruction through the Game-Based Learning (GBL) model assisted by Ladders War media, whereas the control group received Discovery Learning assisted by Question Pull media. Data were analyzed anonymously to maintain student confidentiality.

The study was conducted over four 90-minute sessions (2 × 45 minutes), comprising a pretest, three treatment sessions, and a posttest. Both groups used teaching modules and student worksheets. Before each session, the researcher reviewed the lesson plans and coordinated with the classroom teacher to ensure consistent implementation, while conducting direct observations throughout the intervention. However, the researcher's direct involvement in teaching both groups may have introduced instructor-related bias, a methodological limitation. The relatively short intervention provided focused practice in mathematical critical thinking through structured problem solving, discussion, decision-making, and explanation. Thus, the findings reflect changes in test performance following the intervention rather than long-term development of critical thinking skills. Future research should examine longer interventions and involve independent observers or co-instructors to strengthen instructional fidelity.

In the experimental group, instruction followed six GBL stages: game selection, concept explanation, rule agreement, game implementation, knowledge summarisation, and reflection. Learning was supported by Ladders War, in which students were divided into six groups and solved number-pattern problems using step cards, inter-group battles, ladder boxes, and special boxes containing higher-order-thinking questions. The group that reached the finish line first earned the highest game score.

The control group followed six Discovery Learning stages: stimulation, problem statement, data collection, data processing, verification, and generalization. Learning was supported by Question Pull media, through which students selected number-pattern cards and collaboratively discussed solutions.

Although the intervention period was relatively brief, each treatment session focused on a distinct mathematical critical thinking indicator and incorporated repeated practice through the Ladders War game. The first treatment session focused on the analysis and inference indicators, the second on evaluation and explanation, while the third integrated all four indicators. In each session, the Ladders War game was consistently employed with number pattern tasks adjusted to align with the indicator being addressed. Prior research has demonstrated that game-based learning can support higher-order thinking skills and mathematics learning outcomes (Plass et al., 2015; Mao et al., 2022; Anggoro et al., 2024).

Data were collected using a mathematical critical thinking essay test developed based on four indicators from Facione (2011): analysis, inference, evaluation, and explanation. Four of Facione's six indicators were selected because they were considered appropriate for written assessment of number-pattern material at the elementary school level. The final instrument consisted of five items: one analysis item, one inference item, two evaluation items, and one explanation item.

Content validity was assessed through expert judgment by two validators: a mathematics education lecturer and a practicing fourth-grade elementary school teacher. Each item was evaluated for content relevance, construct alignment, and language clarity using a five-point rating scale. Scores of 4 and 5 were classified as relevant. The results showed an I-CVI of 1.00 for all items and an S-CVI/Ave of 1.00, indicating excellent content validity (Polit & Beck, 2006). Minor revisions to item wording were made based on the validators' feedback.

The instrument was then piloted with students with characteristics comparable to those of the study participants. Item validity was analyzed using corrected item-total correlations, and reliability was assessed using Anates Essay Version 4.0.5. The five items showed item-total correlation coefficients ranging from 0.594 to 0.743, while the reliability coefficient was 0.74, indicating adequate reliability (Guilford, 1956). Thus, all five items were retained as the research instrument.

Students' responses were scored using a predetermined analytic rubric. The researcher scored all responses, while a second rater independently assessed 30 answer sheets. Inter-rater reliability was examined separately for pretest and posttest scores using a two-way random-effects model with absolute agreement. The results indicated excellent reliability for both pretest (single-measure ICC = 0.974, 95% CI [0.941, 0.988]) and posttest scores (single-measure ICC = 0.995, 95% CI [0.989, 0.998]), with both values exceeding the 0.90 threshold (Koo & Li, 2016).

Table 1. Mathematical critical thinking ability test instrument

Number	Capability Indicators Critical Thinking	Questions
1	Students' ability to analyze number patterns in a series of numbers (Analysis)	Notice the following pattern: 20, 16, 12, 8. Each number transitions into the next. Explain the change and determine the difference!

2	Students' ability to make logical conclusions when identifying the continuation of a number sequence (Inference)	Notice the following pattern! 15, 13, 11, 9, … Determine the next number in the pattern and explain how you found the change in each number!
3	Students' ability to judge the truth of a statement grounded in a given number pattern (Evaluation)	Notice the following patterns: 6, 12, 18, 24, ... In Nai's opinion, the next number is 30, while Mel thinks it is 28. Whose opinion is correct? Explain your reasons based on the pattern that is formed!
4	Students' ability to judge the truth of a statement grounded in a given number pattern (Evaluation)	Notice the following pattern: 35, 28, 21, 14, ... According to Dika, the next number is 9, while according to Lina, it is 7. Whose opinion is correct? Explain your reasons based on the pattern that is formed!
5	Students' ability to explain and arrange number patterns based on certain rules (Explanation)	Make a row of 4 numbers with a pattern of decreasing by 9 each Step. It should not be the same as the questions above! Write down the row of numbers you made: (..., ..., ..., ..., ...) After that, please tell me how you can identify the pattern in the numbers you wrote.

Supporting data were gathered through classroom observation and documentation. Student responses were scored using an analytic rubric ranging from 0 to 4, based on answer accuracy, completeness of solution procedures, and clarity of reasoning, as shown in the table. The rubric was applied consistently to minimize scoring subjectivity, and the complete scoring criteria are presented in the table below.

All analyses were conducted using IBM SPSS Statistics Version 30 at an alpha level of 0.05. We first computed descriptive statistics to characterize score distributions across both groups. Before ANCOVA, we examined normality and homogeneity of variance using the Shapiro-Wilk and Levene's tests, respectively. ANCOVA was then performed with pretest scores as the covariate, posttest scores as the dependent variable, and instructional group as the independent variable to account for pre-existing differences in prior ability between groups (Lüdtke & Robitzsch, 2025). The magnitude of the treatment effect was reported using partial eta squared (ç²p) and converted to Cohen's f to interpret effect size based on established criteria (Cohen, 1988). Additionally, a 95% confidence interval was reported to support the interpretation

Table 2. Rubric scoring mathematical critical thinking skills

Answer Criteria	Score
The answer is correct, complete, and accompanied by a reason or an appropriate and systematic completion step	4
The answer is correct, but the completion steps or reasons given are incomplete	3
The response demonstrates partial correctness, although errors remain in the completion process	2
The answer is not correct and shows a low understanding of the concept	1
Not answering or inappropriate answers	0

of findings (Lakens, 2013). As a supplementary measure, normalized gain (N-Gain) was calculated to quantify the degree of improvement in mathematical critical-thinking skills within each group (Hake, 1998).

■ RESULT AND DISCUSSION

Descriptive statistics were used to characterize students' mathematical critical-thinking performance before and after the intervention. As presented in Table 3, the experimental group obtained a higher posttest mean (76.80, SD = 13.14) than the control group (59.60, SD = 16.83). The experimental group's mean increased from 45.40 in the pretest to 76.80 in the posttest, whereas the control group's mean increased from 38.00 to 59.60.

Figure 1 displays the distribution of individual pretest and posttest scores for students in both the experimental and control groups. The diagonal line (y = x) represents the condition in which pretest and posttest scores are equal, allowing the direction and magnitude of each student's score change, as well as any potential ceiling effect, to be observed visually.

Table 3. Descriptive statistical results of pretest and posttest

Group	Data	Score		Mean	Std. Deviation
		Min	Max		
Experiments	Pretest	20	75	45.40	17.436
	Posttest	50	95	76.80	13.140
Controls	Pretest	10	65	38.00	13.769
	Posttest	30	90	59.60	16.828

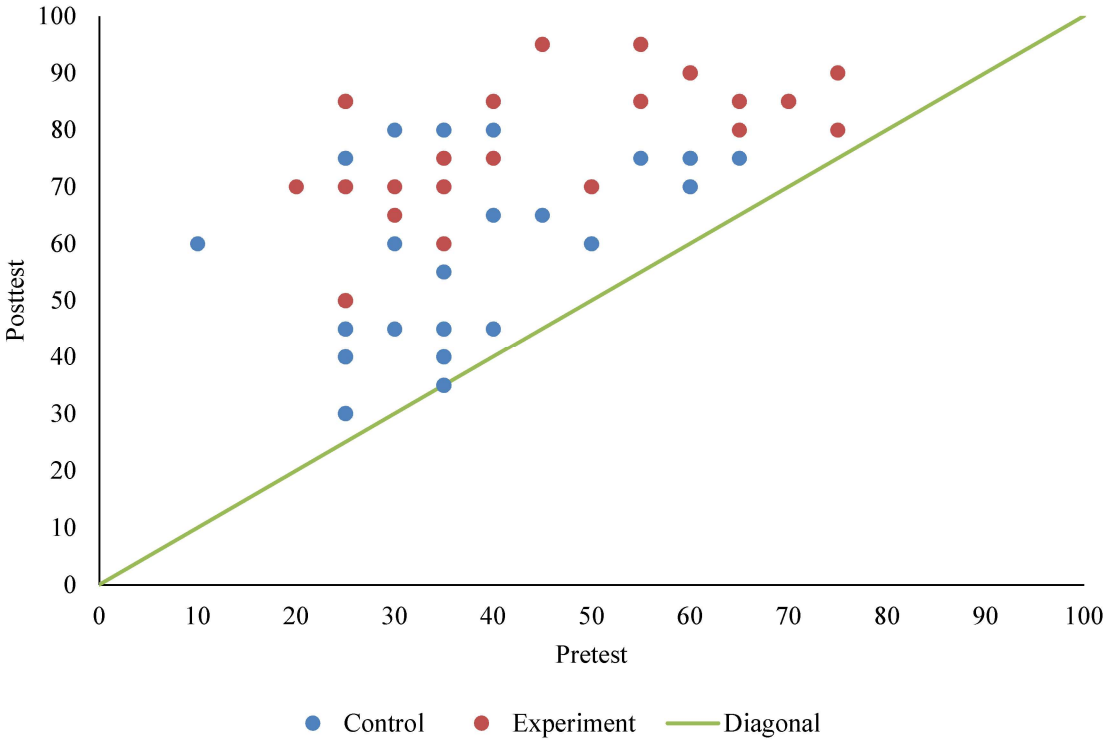


Figure 1. Pretest and posttest scores by group

Figure 1 presents the pretest and posttest scores of individual students in the control group (blue) and the experimental group (orange). The diagonal line marks equal pretest and posttest scores, so points above it indicate improvement. Most points in both groups lie above the diagonal, indicating that most students improved, and the experimental group's points tend to sit higher than those of the control group.

To complement the ANCOVA results, N-Gain analysis was employed descriptively to quantify the magnitude of improvement in mathematical critical thinking skills within each group. Before conducting ANCOVA, the assumptions of normality and homogeneity of variance were examined using the Shapiro-Wilk test, Levene's test, and homogeneity of regression slopes. The results are presented in Table 4.

Table 4. Results of normality, homogeneity of variance, and homogeneity of regression slopes tests

Data	Experimental Sig.	Control Sig.	F	df1	df2	Sig.
Pretest (Shapiro-Wilk)	0.087	0.200	-	-	-	-
Posttest (Shapiro-Wilk)	0.065	0.116	-	-	-	-
Pretest (Levene's test)	-	-	3.808	1	48	0.057
Posttest (Levene's test)	-	-	1.988	1	48	0.165
Pretest × Group (regression-slope homogeneity)	-	-	0.414	1	46	0.523

The assumption tests indicated that the data met the assumptions of normality and homogeneity of variance. The homogeneity of regression slopes assumption was also met, as the interaction between pretest scores and group membership was not statistically significant, $F(1, 46) = 0.414$, $p = 0.523$. Accordingly, ANCOVA was conducted to examine whether students' posttest mathematical critical thinking scores differed significantly between the experimental and control groups after controlling for pretest scores.

Table 5. Results of the analysis of covariance (ANCOVA)

Source	Type III SS	df	MS	F	Sig.	Partial η^2
Pretest	3579.786	1	3579.786	22.859	<0.001	0.327
Classes	2037.883	1	2037.883	13.013	<0.001	0.217
Error	7360.214	47	156.600	-	-	-

The adjusted means further clarify the direction of this effect. As shown in Table 6, the experimental group achieved an adjusted mean of 74.77 (SE = 2.54, 95% CI [69.66, 79.87]), compared with 61.63 (SE = 2.54, 95% CI [56.53, 66.74]) for the control group. The adjusted mean difference was 13.13 points (95% CI [5.81, 20.46]). These results indicate that students who received GBL assisted by Ladders War demonstrated higher mathematical critical-thinking performance than those who received Discovery Learning assisted by Question Pull, even after statistically controlling for differences in prior ability.

Table 6. Comparison of adjusted means

Group	Adjusted Mean	Standard Error	95% CI Lower	95% CI Upper
Experiments	74.766	2.539	69.659	79.873
Controls	61.634	2.539	56.527	66.741

The findings suggest that the experimental condition's effectiveness may be linked to the characteristics of Ladders War, which required students to solve mathematical problems as part of the game's progression. Rather than merely providing an entertaining activity, the game embedded number-pattern problems, inter-group battles, and decision-making into the learning process. These features required students to analyze information, discuss possible solutions, evaluate answers, and explain their reasoning during gameplay. These activities align with the principles of Game-Based Learning, which integrates meaningful game elements with learning objectives to promote active cognitive engagement (Anggoro et al., 2024; Cheng et al., 2023).

From a constructivist perspective, collaborative problem solving during gameplay gave students opportunities to build mathematical understanding through interaction with peers and the teacher (Vygotsky, 1978). The competitive elements of Ladders War also challenged students and gave them opportunities to demonstrate their understanding while working collaboratively. These characteristics are consistent with previous research showing that Game-Based Learning can support students' mathematical thinking through active participation, interaction, n problem-solving (Putra et al., 2024; Rohman & Khaliza, 2025).

To complement the ANCOVA findings, N-Gain analysis was conducted to describe the magnitude of improvement within each group. As shown in Table 7, the experimental group obtained a mean N-Gain of 0.56, whereas the control group obtained 0.33. Both values were classified as moderate according to Hake's (1998) criteria; however, the experimental group demonstrated a greater learning gain.

Table 7. Overall and indicator-level n-gain in mathematical critical thinking

Group	Overall N-Gain	Analysis	Inference	Evaluation	Explanation
Experiments	0.56	0.60	0.57	0.56	0.52
Controls	0.33	0.26	0.22	0.52	0.22

The higher N-Gain in the experimental group complements the ANCOVA result by describing greater improvement relative to students' initial performance. This finding aligns with prior evidence that game-based learning can support meaningful cognitive learning outcomes by increasing student engagement and participation (Barz et al., 2024; Debrenti, 2024). However, the N-Gain result should be interpreted as a descriptive measure of improvement rather than as an additional test of the treatment effect. The ANCOVA provides inferential evidence of differences between groups. The relatively high gain in the analysis indicator in the experimental group may reflect repeated exposure to number-pattern problems during Ladders War. Students had to identify relationships, recognize patterns, compare possible solutions, and choose appropriate strategies before progressing in the game. The activity's collaborative nature also encouraged students to discuss and justify their reasoning, supporting the development of

analytical thinking. This interpretation is consistent with research suggesting that interactive game-based environments can promote discussion and deeper engagement with mathematical problems (Cezarotto et al., 2024; Patterson, 2024).

In contrast, explanation showed the lowest N-Gain among the four indicators in the experimental group. This suggests that although gameplay supported students in analyzing and solving problems, systematically expressing mathematical reasoning in written form remained challenging. Oral discussion during gameplay does not necessarily translate directly into organized written explanations. Therefore, future implementation could complement gameplay with structured written reflection or justification activities to give students more opportunities to practice communicating their reasoning (Christopoulos et al., 2024).

The relatively small difference between the two groups on the evaluation indicator warrants cautious interpretation. This indicator was measured by two items (items 3 and 4), accounting for 40% of the total test, whereas each of the remaining three indicators was assessed by a single item. This disproportionate weighting may have influenced the N-Gain value for the evaluation indicator, as the control group obtained a notably higher gain on this indicator (0.52) compared with the other three indicators (0.22–0.26).

However, since the difficulty level and scoring patterns of individual items were not analyzed separately, and since the analytic rubric was applied uniformly across all items without item-level calibration, the specific factors contributing to the control group's relatively high N-Gain on this indicator, whether attributable to lower item difficulty, scoring rubric characteristics, or other influences, cannot be determined with certainty. Accordingly, this finding is interpreted descriptively and is not used to conclude that either instructional model is more effective in developing evaluation skills.

The findings provide empirical evidence that a non-digital, strategy-based game can support the development of mathematical critical thinking when game mechanics are directly integrated with higher-order mathematical tasks. In particular, Ladders War combines collaborative problem-solving, competition, decision-making, and mathematical challenges within the Game-Based Learning framework. This extends existing evidence on Game-Based Learning by demonstrating its application to mathematical critical thinking in elementary education and within number-pattern learning.

Practically, Ladders War offers teachers an alternative instructional medium for engaging elementary students in mathematical problem solving without requiring digital infrastructure. Teachers can adapt its mechanics to other mathematical topics that require analysis, reasoning, evaluation, and explanation. Nevertheless, because the explanation indicator showed the smallest gain, gameplay should be complemented by structured written reflection to strengthen students' ability to communicate mathematical reasoning.

The findings should be interpreted in light of several limitations. The study involved only 50 students from one elementary school and focused on number-pattern material over four sessions, which limits the generalisability of the findings. Future studies should involve larger, more diverse samples, examine other mathematical topics, and use longer interventions.

■ CONCLUSION

The Game-Based Learning model assisted by Ladders War was associated with higher mathematical critical thinking performance among elementary school students in number-pattern learning after controlling for pretest scores, with an adjusted mean difference of 13.13 points. Ladders War integrated problem solving, discussion, decision-making, and mathematical explanation into game progression, providing

structured opportunities to practice critical thinking. As a non-digital instructional medium, it can be integrated into elementary mathematics learning and complemented with structured reflection and written justification. However, interpret the findings with consideration of the short intervention, researcher involvement, and single-school sample. Future research should involve longer interventions, larger and more diverse samples, and independent instructors or observers.

■ DECLARATION OF GENERATIVE AI USAGE IN THE WRITING PROCESS

During the preparation of this manuscript, the authors used Claude AI to help refine sentence structure and translate the text. Following the use of this tool, the authors reviewed and revised the content as necessary and accept full responsibility for the final content of the article.

■ REFERENCES

Alotaibi, M. S. (2024). Game-based learning in early childhood education: A systematic review and meta-analysis. *Frontiers in Psychology*, *15*. https://doi.org/10.3389/fpsyg.2024.1307881

Anggoro, B. S., Dewantara, A. H., Suherman, S., Muhammad, R. R., & Saraswati, S. (2024). Effect of game-based learning on students' mathematics higher-order thinking skills: A meta-analysis. *Revista de Psicodidáctica*, *30*(1), 500158. https://doi.org/10.1016/j.psicoe.2024.500158

Azizah, F. K., & Fidrayani. (2024). The influence of the use of educational game learning media "Snake and Ladder" on mathematics learning outcomes elementary school students. *At-Taqaddum*, *16*(2), 93–107. https://journal.walisongo.ac.id/index.php/attaqaddum/article/view/21896

Ballance, O. J. (2024). Sampling and randomisation in experimental and quasi-experimental CALL studies: Issues and recommendations for design, reporting, review, and interpretation. *ReCALL*, *36*(1), 58–71. https://doi.org/10.1017/S0958344023000162

Barz, N., Benick, M., Dörrenbächer-Ulrich, L., & Perels, F. (2024). The effect of digital game-based learning interventions on cognitive, metacognitive, and affective-motivational learning outcomes in school: A meta-analysis. *Review of Educational Research*, *94*(2), 193–227. https://doi.org/10.3102/00346543231167795

Cezarotto, M. A., Martinez, P. N., Torres Castillo, R. C., Stanford, T., Engledowl, C., Degardin, G., & Chamberlin, B. (2024). Open-ended mathematics learning: Implications from the design of a sandbox game. *International Journal of Game-Based Learning*, *14*(1), 1–19. https://doi.org/10.4018/IJGBL.337795

Cheng, Y., Lai, C., Chen, Y., Wang, W., Huang, Y., & Wu, T. (2023). Enhancing student's computational thinking skills with student-generated questions strategy in a game-based learning platform. *Computers & Education*, *200*, 104794. https://doi.org/10.1016/j.compedu.2023.104794

Christopoulos, A., Mystakidis, S., Kurczaba, J., Laakso, M.-J., & Stylios, C. (2024). Is immersion in 3D virtual games associated with mathematical ability improvement in game-based learning? *International Journal of Science and Mathematics Education*, *22*(7), 1479–1499. https://doi.org/10.1007/s10763-023-10440-4

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Dan, N. N., Trung, L. T. B. T., Nga, N. T., & Dung, T. M. (2024). Digital game-based

learning in mathematics education at primary school level: A systematic literature review. *Eurasia Journal of Mathematics, Science and Technology Education*, *20*(4), em2423. https://doi.org/10.29333/ejmste/14377

Debrenti, E. (2024). Game-based learning experiences in primary mathematics education. *Frontiers in Education*, *9*, 1–8. https://doi.org/10.3389/feduc.2024.1331312

Dienes, Z. P. (1963). On the learning of mathematics. *The Arithmetic Teacher*, *10*(3), 115–126. https://doi.org/10.5951/AT.10.3.0115

Facione, P. A. (2011). *Critical thinking: What it is and why it counts*. Insight Assessment. https://www.insightassessment.com/CT-Resources/Teaching-For-and-About-Critical-Thinking/Critical-Thinking-What-It-Is-and-Why-It-Counts/Critical-Thinking-What-It-Is-and-Why-It-Counts-PDF

Guilford, J. P. (1956). *Fundamental statistics in psychology and education* (3rd ed.). McGraw-Hill.

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

Halidah, F., Aini, I. I. N., Ernawati, N., Juhaeni, J., Faizah, L., & Safaruddin, S. (2024). *Pengembangan media pembelajaran matematika berbasis permainan ular tangga untuk meningkatkan hasil belajar siswa* [Development of mathematics learning media based on the snakes and ladders game to improve student learning outcomes]. *Journal of Instructional and Development Researches*, *4*(2), 64–74. https://doi.org/10.53621/jider.v4i2.306

IEA. (2023). *TIMSS 2023 international results in mathematics and science*. https://timss2023.org/results/

Jannah, R., Soraya, R. A., Suriansyah, A., & Cinantya, C. (2024). *Kemampuan berpikir kritis dalam pembelajaran matematika terhadap hasil belajar di sekolah dasar* [Critical thinking skills in mathematics learning on learning outcomes in elementary schools]. *MARAS: Jurnal Penelitian Multidisiplin*, *2*(4), 1991–1998. https://doi.org/10.60126/maras.v2i4.550

Klooger, M. (2026). More than fun and games: Teaching mathematics through games. *Mathematics Teacher Education and Development*, *28*(1). https://mted.merga.net.au/index.php/mted/article/view/1348

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863. https://doi.org/10.3389/fpsyg.2013.00863

Lüdtke, O., & Robitzsch, A. (2025). ANCOVA versus change score for the analysis of two-wave data. *The Journal of Experimental Education*, *93*(2), 363–395. https://doi.org/10.1080/00220973.2023.2246187

Mao, W., Cui, Y., Chiu, M. M., & Lei, H. (2022). Effects of game-based learning on students' critical thinking: A meta-analysis. *Journal of Educational Computing Research*, *59*(8), 1682–1708. https://doi.org/10.1177/07356331221100740

Maula, N. F., Suryadi, D., & Jupri, A. (2024). A systematic literature review: Generalization in solving number pattern problems. *Jurnal Pendidikan MIPA*, *25*(2), 701–711. https://doi.org/10.23960/jpmipa/v25i2.pp701-711

Monteleone, C., Miller, J., & Warren, E. (2023). Conceptualising critical mathematical thinking in young students. *Mathematics Education Research Journal*, *35*, 339–359. https://doi.org/10.1007/s13394-023-00445-1

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*.

Patterson, J. M. (2024). Repurposing video games as discussion tools. *International Journal of Game-Based Learning*, *14*(1), 1–19. https://doi.org/10.4018/IJGBL.349589

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist*, *50*(4), 258–283. https://doi.org/10.1080/00461520.2015.1122533

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, *29*(5), 489–497. https://doi.org/10.1002/nur.20147

Pramasdyahsari, A. S., Setyawati, R. D., Aini, S. N., Nusuki, U., Arum, J. P., Astutik, I. D., Widodo, W., Zuliah, N., & Salmah, U. (2023). Fostering students' mathematical critical thinking skills on number patterns through digital book STEM PjBL. *Eurasia Journal of Mathematics, Science and Technology Education*, *19*(7). https://doi.org/10.29333/ejmste/13342

Putra, L. D., Arlinsyah, N. D., Ridho, F. R., Syafiqa, A. N., & Annisa, K. (2024). *Pemanfaatan Wordwall pada model game based learning terhadap digitalisasi pendidikan sekolah dasar* [Utilization of Wordwall in game-based learning models for the digitalization of elementary school education]. *Jurnal Dimensi Pendidikan dan Pembelajaran*, *12*(1), 81–95. https://doi.org/10.24269/dpp.v12i1.8894

Rohman, A. D., & Khaliza, T. N. (2025). *Inovasi pembelajaran abad ke-21 melalui media Baamboozle: Upaya meningkatkan keterampilan berpikir kritis dan kolaboratif pada peserta didik sekolah dasar* [21st century learning innovation through Baamboozle media: Efforts to improve critical and collaborative thinking skills in elementary school students]. *Journal of Elementary Education Research*, *1*(2), 98–106. https://doi.org/10.70162/jeer.v1i2.38

Sukmawati, S., & Arifanti, D. R. (2022). *Perbandingan antara minat belajar matematika siswa dalam kelas yang berjumlah besar dengan berjumlah kecil* [Comparison between students' interest in learning mathematics in large and small classes]. *Pedagogy: Jurnal Pendidikan Matematika*, *7*(1), 70–81. https://doi.org/10.30605/pedagogy.v7i1.1802

Thornhill-Miller, B., Camarda, A., Mercier, M., Burkhardt, J.-M., Morisseau, T., Bourgeois-Bougrine, S., Vinchon, F., El Hayek, S., Augereau-Landais, M., Mourey, F., Feybesse, C., Sundquist, D., & Lubart, T. (2023). Creativity, critical thinking, communication, and collaboration: Assessment, certification, and promotion of 21st century skills for the future of work and education. *Journal of Intelligence*, *11*(3). https://doi.org/10.3390/jintelligence11030054

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. https://doi.org/10.2307/j.ctvjf9vz4

Wardani, M. S., & Himmah, W. I. (2026). Traditional game-based learning for mathematics in elementary school: A systematic literature review. *Desimal: Jurnal Matematika*, *9*(2), 228–254. https://doi.org/10.24042/djm.v9i2.31699

Widiastika, N., Kartika, H., & Hanifah, H. (2023). *Pengaruh model pembelajaran CORE terhadap kemampuan koneksi matematis siswa: Studi quasi eksperimen* [The effect of the CORE learning model on students' mathematical connection abilities: A quasi-experimental study]. *Journal of Research in Science and Mathematics Education*, *2*(3), 173–185. https://doi.org/10.56855/jrsme.v2i3.490