Jurnal Pendidikan MIPA
e-ISSN: 2685-5488 | p-ISSN: 1411-2531
URL: https://jpmipa.fkip.unila.ac.id/index.php/jpmipa
Jounal's DOI: http://doi.org/10.23960/jpmipa

Vol. 27, No. 3, 2026
Article ID: 1434
Pages: 1538-1554

The Role of Zep Quiz Gamification in Improving Junior High School Students' Mathematical Problem-Solving Abilities

Gusti Peanti[1,*], Agus Susanta[1], Effie Efrida Muchlis[1], Hanifah, & Siti Rahaimah Ali[2]
[1]Department of Magister Mathematics Education, Universitas Bengkulu, Indonesia
[2]Department of Mathematics Education, Universiti Pendidikan Sultan Idris, Malaysia

*Corresponding email: gustipeanti@graduate.unib.ac.id

Received: 15 June 2026 *Accepted: 13 August 2026* *Published: 29 August 2026*

Abstract: Mathematical problem-solving ability is a crucial competency, yet it remains a challenge in many schools, including SMPIT Al Qalam South Bengkulu, where students' statistics scores fall into the "very low" category (48.6%) due to conventional teaching methods that lack significant technological integration. This study aimed to examine the effectiveness of the Zep Quiz 2D virtual world-based gamification platform on 8th-grade students' mathematical problem-solving abilities in statistics. A quasi-experimental pretest-posttest control group design was employed, involving 60 students. Samples were selected via purposive sampling: 30 students in the experimental class (VIII C) were taught using Zep Quiz integrated with Polya-based student worksheets, while 30 students in the control class (VIII D) received conventional instruction. The research instrument consisted of a descriptive test validated and proven reliable based on Polya's four problem-solving indicators. Data were analyzed using descriptive statistics and an independent t-test after verifying the assumptions of normality (Shapiro-Wilk test) and homogeneity (Levene's test). The results showed that the experimental class achieved a higher average posttest score of 87.30 (N-gain of 0.77, high category) compared to the control class, which had an average posttest score of 81.93 (N-gain of 0.66, medium category); the hypothesis test yielded a significant result ($p = 0.038 < 0.05$). It is concluded that Zep Quiz is significantly more effective than conventional instruction in enhancing students' mathematical problem-solving abilities.

Keywords: gamification, mathematical problem solving, quasi-experiment, zep quiz.

Article's DOI: https://doi.org/10.23960/jpmipa.v27i3.pp1538-1554

INTRODUCTION

Mathematical problem-solving ability is a core competency explicitly mandated in the national mathematics learning objectives based on the Minister of National Education Regulation, including problem understanding, model design, solution, and solution interpretation (Sriwahyuni & Maryati, 2022). Riyanto and Amidi (2024) emphasized that problem-solving is at the heart of mathematics because it requires simultaneous mastery of the material, solution strategies, self-monitoring, and a productive attitude. This ability directly influences learning motivation (Liu et al., 2025) and the quality of problem solving (Kopparla et al., 2019; Wilkie, 2024; Rocha & Babo, 2024), and is influenced by demographic background (Amalina & Vidákovich, 2023), where prior knowledge of mathematics is a dominant factor (Dewi et al., 2025), and gender differences in responding to teacher instructions (Schreiber & Ashkenazi, 2024). Meanwhile, As Sa'adah and Wulandari (2022) emphasized that this ability is not just a general goal but also a core process and basic skill in mathematics, providing students with the means to face routine and non-routine problems in real life (Liu et al., 2025). However, the reality on the ground is the opposite: the results of the 2022 PISA placed Indonesia in 69th place out of 81 countries with an average mathematics score of only 366, experiencing a significant decline compared to the 2015–2018 period and lagging far behind five other ASEAN countries (Masito et al., 2024). More than 63% of junior high school students reported that learning in school remains less interesting because it is overly theoretical and lacks interactive activities (Sausan & Wibowo, 2024). This condition illustrates the quality of mathematics learning in Indonesia, which can no longer be addressed with conventional approaches.

Various efforts to improve problem-solving skills have been extensively researched. Among them, learning based on thinking processes and creative problems has been shown to improve students' understanding and emotional engagement (Chaiarwut et al., 2025; Larsson et al., 2024; Bishara, 2016); cooperative models with a contextual approach (CTL) have a positive impact on mathematics success (Hariyati & Purwanto, 2023); lesson study (Bayaga, 2024) and experiential learning (Mayoral-Rodríguez et al., 2018) have been shown to improve problem-solving skills. Meanwhile, blended problem-based learning in flipped online classes sharpens critical thinking and practical skills (Pathuddin et al., 2025). Informal channels have also been explored, ranging from math clubs (Turner et al., 2011) to parent-involved homework (Williams & Williams, 2021). Despite their diversity, these approaches have not yet optimally accommodated the characteristics of the digital generation, who grew up in a technological environment and are more responsive to digital-based learning methods. The Indonesian government, through the 2013 Curriculum, has emphasized the importance of the 4C skills: critical thinking and problem-solving, creativity, communication, and collaboration (Sausan & Wibowo, 2024). 21st-century learning, namely technology-based student-centered learning (Liu et al., 2025), where artificial intelligence-based learning environments have been shown to have a positive and significant impact on mathematical problem-solving skills, with the note that the use of well-designed AI can provide quality guidance as long as teachers remain active in supervising (Cãprioarã, 2015). Effective digital platforms must have an attractive design that does not burden students' cognitive abilities to maintain motivation (Yuliati, 2021; Munggaran et al., 2025; Salamah et al., 2025).

Gamification presents the most targeted solution to address these challenges. Research consistently demonstrates that scoring, badges, and leaderboard elements not only increase engagement but also strengthen intrinsic motivation and improve cognitive learning outcomes by an average of 10–15 points (Aprijal et al., 2020; Mägdefrau et al., 2025; Perche et al., 2025; Holmes et al., 2025; Hasibuan, 2021; Mustafa et al., 2023; Mägdefrau et al., 2025; Fryer et al., 2025; Yeribatuah & Arthur, 2023). This alignment is further strengthened by the characteristics of junior high school students aged 12–15 years. This early adolescence phase is characterized by a high need for social recognition, sensitivity to rewards, and a strong sense of adventure (Mustafa et al., 2023). Gamification is an ideal bridge between play and learning (Ramadhani et al., 2025), and Montero-Izquierdo et al. (2024) emphasized that integrating technology and game elements can change perceptions of subjects considered difficult, making them more interesting and emotionally manageable. One of the most innovative gamification platforms today is Zep Quiz. An EdTech platform that integrates digital quizzes into an interactive 2D virtual world, where students move avatars to find and solve challenging questions on a virtual map. Several previous studies have confirmed the use of Zep Quiz for learning, including an increase in learning motivation from 72% to 87% (Chieng et al., 2021). Meanwhile, Hilmi and Jasiah (2025) confirmed that Zep Quiz increased student interaction and interest, even surpassing Kahoot in a comparative study. Broader research on interactive media also provides a strong foundation. Daryanes et al. (2023) proved that interactive media based on Articulate Storyline is effective in improving problem-solving skills; Ramadhani et al. (2025), through a systematic literature review of 30 articles, concluded that gamification is effective in improving problem-solving skills at elementary and junior high school levels. Sari et al. (2025) showed that problem-based digital teaching materials can develop students' mathematical problem-solving skills.

However, significant research gaps remain. First, almost all studies on Zep Quiz only measure general learning outcomes and motivation. None have specifically tested its effectiveness on mathematical problem-solving skills using structured indicators such as Pólya's stages. Second, gamification at the junior high school level remains far less common than at the elementary, high school, and university levels, even though junior high school students' psychological characteristics are most compatible with gamification mechanisms (Anisah et al., 2024). Third, no research has specifically integrated Zep Quiz into junior high school statistics material, which has been empirically proven to be the most difficult and to require the most engaging, contextual approach. This condition aligns with initial findings at SMPIT Al Qalam Bengkulu Selatan: interviews with eighth-grade mathematics teachers revealed that student scores in statistics were in the very low category (48.6%), most students had difficulty working on statistics problems, and the learning pattern was still entirely conventional, without interactive media. The novelty of this research lies in three aspects that have never been studied simultaneously: (1) testing the effectiveness of Zep Quiz through a controlled experimental design (quasi-experimental pretest-posttest control group design) on mathematical problem-solving abilities as measured by the Polya indicator, (2) emphasizing statistical material, especially data centralization at the junior high school level, and (3) the use of students' worksheets directly integrated with the Zep Quiz platform as a structured learning instrument. Thus, this research not only fills an existing gap but also produces new empirical evidence on the effectiveness of gamification in improving junior high school students' higher-order thinking skills in a context that has been neglected so far, while also offering a replicable and applicable learning model for mathematics teachers in the digital era.

Based on the research gaps and novelty previously outlined, this study aims to determine whether the use of the interactive learning medium "Zep Quiz" is effective in enhancing students' mathematical problem-solving skills regarding statistics, specifically for Grade VIII students at SMPIT Al Qalam, South Bengkulu, guided by the following research questions:

1. What are the mathematical problem-solving abilities of 8th-grade students at SMPIT Al Qalam South Bengkulu regarding the topic of statistics before and after the implementation of the Zep Quiz interactive learning medium?
2. Is the Zep Quiz interactive learning medium effective in improving the mathematical problem-solving abilities of 8th-grade students at SMPIT Al Qalam South Bengkulu regarding the topic of statistics, when compared to conventional instruction?
3. To what extent does the use of the Zep Quiz interactive learning medium influence the improvement of mathematical problem-solving abilities among 8th-grade students at SMPIT Al Qalam South Bengkulu regarding the topic of statistics (measures of central tendency), as measured by the indicators of Pólya's stages?

Based on the formulation of the research question, the hypothesis proposed in this study is:

H0: The Zep Quiz interactive learning medium is ineffective in developing the mathematical problem-solving abilities of 8th-grade students at SMPIT Al Qalam South Bengkulu on the topic of statistics.

Ha: The Zep Quiz interactive learning medium effectively enhances the mathematical problem-solving skills of 8th-grade students at SMPIT Al Qalam South Bengkulu on the topic of statistics.

METHOD

Population and Sample

The study population consisted of all Grade VIII students at SMPIT Al Qalam South Bengkulu during the 2025/2026 academic year, distributed across four study groups. These four classes were homogeneous, possessing relatively equal academic abilities as evidenced by their first-semester final assessment results. Two classes were selected as the research sample: Class VIII C served as the experimental class, and Class VIII D served as the control class. The experimental class received a specific treatment involving the method of material delivery (using instructional media) and the completion of practice exercises via Student Worksheets supported by the Zep Quiz application. All questions within the application corresponded to the C4 cognitive level; students were required to detail their solution processes on the student Worksheet, ensuring that every step of the solution aligned with the four indicators of Polya's problem-solving stages.

Furthermore, the application guided students to ensure that each step matched the procedure outlined in the students' worksheets. In contrast, the control class received conventional instruction, from delivering material (using the lecture method) to completing textbook exercises, which did not exclusively target the C4 cognitive level. Students completed the practice exercises in notebooks. However, while working on the problems, students were still required to solve them using Polya's steps—recalling the teacher's explanations or consulting their notes from the lesson. Consequently, no additional guidance was provided via structured worksheets to facilitate the sequential application of Polya's steps, nor was the Zep Quiz application used. Details regarding the research sample are presented in Table 1.

Table 1. Research sample

	VIII C (Experiment)	VIII D (Control)
	30	30
Total	30	30

Research Design

This study employed a quantitative approach with a quasi-experimental pretest-posttest control group design. The study involves two groups: an experimental class that received the treatment using Zep Quiz media, and a control class that received no additional treatment. The population for this study consists of all Grade VIII students at SMPIT Al Qalam, South Bengkulu. Table 2 presents the research design.

Table 2. Research design

Class	Pretest	Treatment	Postest
Experimental (VIII C)	O_1	X (Zep Quiz)	O_2
Control (VIII D)	O_1	Y (conventional)	O_2

This study obtained official research authorization from SMPIT Al Qalam South Bengkulu via Research Permit Number 19/SMPIT.AQ/SI/K/II/2026. Additionally, the researchers obtained informed consent from the student subjects before the study began. All research procedures adhered to ethical principles, including maintaining the confidentiality of student data and ensuring that participation was voluntary and caused no harm to the subjects.

The research was conducted at SMPIT Al Qalam Bengkulu Selatan, located on Jalan Pemangku Basri, Tanjung Mulia Village, Pasar Manna District, South Bengkulu Regency, Bengkulu Province. The research took place in the even semester of the 2025/2026 academic year, specifically from February to April 2026. The research population comprised all eighth-grade students, organized into four study groups totaling 100 students. Purposive sampling was used. Classes at the IT school are separated into male and female classes. VIII A and VIII B are male classes, while classes VIII C and VIII D are female classes. Classes VIII C and VIII D were chosen because the number of students was appropriate for the research sample, namely, one class of 30 students each. Meanwhile, the male classes consisted of only 20 students each (failing to meet the research sample requirements). Furthermore, the two female classes were academically homogeneous based on their final semester assessment results. The homogeneity test using Levene's Test yielded a Levene statistic of 0.822, with degrees of freedom df1 = 1 and df2 = 58, and a significance value of 0.368. Since the significance value is greater than or equal to 0.05 (e" 0.05), the samples were drawn from a homogeneous population.

Research Instruments

The instrument used in this study was a mathematical problem-solving ability test consisting of two essay questions per testing session. Each question item was designed based on Polya's four problem-solving indicators: (1) understanding the problem; (2) devising a plan; (3) carrying out the plan; and (4) looking back. The essay format was selected because problem-solving ability can be validly observed only through students' written reasoning processes, rather than merely through final answers as in multiple-choice formats.

Student responses were evaluated using an analytic rubric based on Polya's four problem-solving indicators: understanding the problem, devising a plan, carrying out the plan, and reflecting on the outcome. Since each indicator reflects a different level of cognitive complexity, the maximum score for each indicator is weighted accordingly. The "understanding the problem" stage carries a maximum score of 2. Students receive the maximum score of 2 if they completely and correctly state the given information and the unknown (what is being asked), including appropriate notation or units. A score of 1 is awarded if the student states the given information or the unknown but does so incompletely, makes errors regarding notation or units, or includes information that is not relevant to the problem. A score of 0 is assigned if the student fails to state the given information and the unknown entirely, or if the written response is unrelated to the problem provided.

Meanwhile, the devising-a-plan stage has a maximum score of 3. Students receive a score of 3 if they select and write down a solution strategy or method that is appropriate and relevant to the problem, while also demonstrating logical and systematic connections between steps. A score of 2 is awarded if students select a relevant strategy or method but provide an incomplete explanation of the connections between steps or include an inaccurate part in the plan. A score of 1 is given if students write down a plan, but the chosen strategy or method is inappropriate or does not align with the context of the problem. A score of 0 is assigned if students do not write a plan at all, or if the written plan bears no relation to the given problem.

The stage of carrying out the plan has a maximum score of 3. A score of 3 is awarded if the student correctly executes all solution steps according to the plan, meaning calculations or procedures are performed accurately and systematically, resulting in the correct answer. A score of 2 is awarded if the student follows the planned solution steps but makes minor calculation or procedural errors that do not affect the overall solution process. A score of 1 is awarded if the student carries out only some of the solution steps but commits significant procedural or calculation errors, resulting in an incorrect answer. A score of 0 is awarded if the student fails to carry out any solution steps, or if the steps taken do not align with the plan and fail to yield a meaningful answer.

The Looking Back stage has a maximum score of 2. Students receive a score of 2 if they verify the correctness of their answer and its alignment with the question, explain the reasoning behind the answer, and draw a conclusion regarding the solution that is both accurate and complete. A score of 1 is awarded if students review the answer and/or write a conclusion. However, the verification is incomplete, contains minor errors, or the written conclusion does not fully align with the obtained solution. A score of 0 is given if students do not review the answer or write a conclusion, or if the written conclusion is unrelated to the problem being solved.

The maximum score a student can obtain for each essay question is 10, derived from the components (2 + 3 + 3 + 2); thus, for two essay questions, the maximum raw score is 20. This total raw score is then converted to a 0–100 scale using the formula: Final Score = (Total Score × 10) / 2, to enable parametric statistical analysis.

Before use in the study, all instruments, including the teaching module, student worksheets, Zep Quiz learning media, and the problem-solving ability test questions, were validated by four experts who met the academic qualifications for validators. Instrument validity was calculated using Aiken's formula, with a validity criterion of V e" 0.70, while reliability was calculated using Hoyt's ANOVA, with a reliability criterion of r11 e" 0.70. In addition to expert validation and reliability testing, the problem-solving ability test questions were piloted with a class outside the research sample to determine the difficulty level and discriminating power of the questions.

Validation results indicated that the teaching module for the experimental class achieved a validity score of 0.83 and a reliability score of 0.73, while the teaching module for the control class achieved a validity score of 0.82 and a reliability score of 0.82; both sets of values demonstrate high levels of validity and reliability. The Zep Quiz learning medium obtained a validity score of 0.83 (highly valid) and a reliability score of 0.80 (highly reliable). The students' worksheets achieved a validity score of 0.81 and a reliability score of 0.75, also classifying them as highly valid and reliable. Meanwhile, the problem-solving ability test instrument obtained validity scores of 0.89 for question number 1 and 0.87 for question number 2, with a reliability score of 0.75, indicating reliability. In summary, Table 3 presents the validation results.

Table 3. Instrument validation by the expert team/panel

No	Aspect	Validity	Reliability(r_{11})	description
1	Zep Quiz-Assisted Teaching Module	0.83	0.73	Valid and Reliable
2	Conventional Teaching Module	0.82	0.82	Valid and Reliable
3	Zep Quiz Learning Media	0.83	0.80	Valid and Reliable
4	Student Worksheet	0.81	0.75	Valid and Reliable

Meanwhile, Table 4 presents the results of the expert team's validation of the problem-solving ability instrument.

Table 4. Validation of the mathematical problem-solving ability test instrument by the expert team/ panel

No	Question Level / Question Form	Validity	Reliability (r_{11})	description
1	Level C4 / Essay	0.89	0.75	Validity and Reliability
2	Level C4 / Essay	0.87		Validity and Reliability

Data Analysis Procedure

Data analysis was conducted in stages, with each statistical test selected based on meeting assumptions rather than by default. Descriptive analysis utilized N-Gain scores rather than simply the difference between post-test and pre-test scores. Before hypothesis testing, three parametric statistical assumptions were examined using SPSS, namely: the normality test (Shapiro-Wilk), homogeneity test (Levene), and linearity test (linear regression), with the criterion that data satisfy the assumptions if the Sig. value is e" 0.05. These three tests serve as prerequisites for further data processing. If any assumption is not met, the data cannot be analyzed using parametric statistics; the alternative is to use a non-parametric test, specifically the Mann-Whitney test. Since all assumptions for the research data were met, namely that the data were normally distributed, homogeneous, and linear, hypothesis testing was conducted using the Independent-Samples T-Test (á = 0.05) on the N-Gain data, with the following statistical hypotheses:

H_0 : µ1 = µ2 (There is no significant difference)
H_1 : µ1 “‘ µ2 (There is a significant difference.)
Testing criteria: H€ is rejected if the Sig. value. (2-tailed) < 0,05.

Using N-Gain scores, rather than post-test scores alone, helps ensure that statistically detected differences reflect the magnitude of improvement from the treatment, rather than residual differences in initial ability between groups.

■ RESULT AND DISCUSSION

Description of Pretest and Posttest Results Data

Based on the administration of pretests and posttests in both classes, descriptive statistical data were obtained, as presented in Figure 1 below.

The box plot displays the interquartile range from the first to the third quartile, with the central line representing the median; the whiskers indicate the minimum and maximum values (excluding outliers), the Ç%symbol represents the mean, and black dots denote outliers. N = 30 for each group. Figure 1 presents a comparison of pretest and posttest score distributions between the experimental and control groups using paired box plots. The pretest results show relatively comparable data characteristics across both groups, with medians hovering around 50 and means of 45.5 for the experimental group and 45.3 for the control group. This similarity indicates that the two groups had homogeneous baseline mathematical problem-solving abilities before the intervention.

Following the instruction, post-test results showed score improvements in both groups, albeit

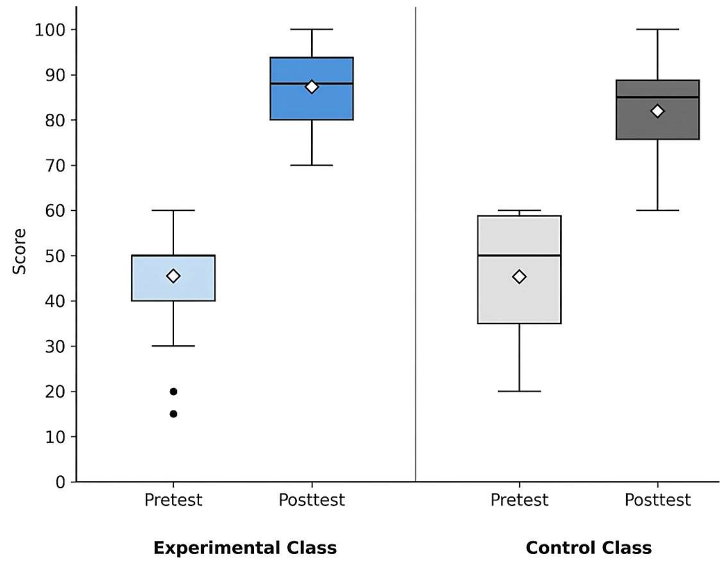

Figure 1. Comparison of pretest and posttest scores in the experimental and control classes

to varying degrees. The experimental class achieved a median of 88 and a mean of 87.3, whereas the control class recorded a median of 85 and a mean of 81.9. Furthermore, while the interquartile range (IQR) for the experimental class (80–94) was relatively comparable to that of the control class (76–89), the upper-bound values and mean scores for the experimental class were consistently higher than those of the control class. This indicates that the distribution of student scores in the experimental class was more concentrated toward the maximum score of 100. Additionally, two outlier scores of 15 and 20 were observed in the experimental class's pre-test results, indicating a small number of students with initial proficiency levels far below the group average. These findings initially indicate that implementing a gamification-based learning model (Zep Quiz) in the experimental class was more effective at improving students' mathematical problem-solving skills than the conventional instruction used in the control class.

Based on Figure 1 above, a clear difference in effectiveness is evident between the experimental and control classes. Visually, the experimental class shows a more robust and consistent progression toward the maximum score of 100 than the control class. Furthermore, this pattern provides a strong preliminary indication that the gamification-based learning model (Zep Quiz) implemented in the experimental class yields more optimal acceleration of learning, particularly in students' mathematical problem-solving abilities.

Prerequisite Analysis Test

Before hypothesis testing, the fulfillment of parametric statistical assumptions was verified through a series of prerequisite tests.

Normality test

Normality testing was conducted using the Shapiro-Wilk method in SPSS, with the criterion that data are considered normally distributed if the significance value is greater than 0.05. The test results showed significance values of 0.114 for the experimental group's post-test and 0.156 for the control group's post-test. Both values exceeded 0.05; therefore, the post-test data for both the experimental and control groups were considered normally distributed.

Homogeneity test

Homogeneity testing was conducted using Levene's test, with the criterion that data are considered homogeneous if the significance value is greater than 0.05. The test results yielded a post-test significance value of 0.408, which exceeds the significance level of á = 0.05. Consequently, the null hypothesis (H€) is accepted. It can thus be concluded that the variances of the post-test data for the experimental and control classes are significantly homogeneous.

Hypothesis Testing

Once all parametric assumptions were met, hypothesis testing was conducted using an independent-samples t-test at a significance level of á = 0.05. The test results yielded a significance value (2-tailed) of 0.038. Since this value is less than 0.05 (0.038 < 0.05), the null hypothesis (H€) was rejected. Therefore, there is a statistically significant difference in mathematical problem-solving ability in statistics between students who used Zep Quiz and those who received conventional instruction.

N-Gain Score Analysis

N-gain calculations were conducted to assess the treatment's relative effectiveness in improving mathematical problem-solving abilities in each class. Overall, the experimental class achieved a mean pretest score of 45.50 and a mean posttest score of 87.30, with an N-Gain of 0.77, placing it in the high category. Meanwhile, the control class obtained a mean pretest score of 45.18 and a mean posttest score of 81.93, with an N-Gain of 0.66, placing it in the medium category. Details regarding the N-Gain for each indicator of mathematical problem-solving ability (Polya's stages) are presented in Table 5.

Table 5. N-gain score for each mathematical problem-solving ability indicator

Polya Stage	Pretest (Exp.)	Posttest (Exp.)	N-Gain (Exp.)	Criteria (Exp.)	Pretest (ctrl.)	Posttest (ctrl.)	N-Gain (ctrl.)	Criteria (ctrl.)
Understanding the Problem	2.40	3.63	0.77	High	2.77	3.63	0.70	High
Resolution Plan	2.97	5.03	0.68	Medium	2.93	4.90	0.64	Medium
Implementing the Resolution Plan	2.80	5.83	0.95	High	2.60	5.27	0.78	High
crosscheck	0.93	3.00	0.67	Medium	0.77	2.60	0.57	Medium

Treatment effectiveness is measured using an N-gain score calculated as the difference between the posttest and pretest scores, divided by the difference between the ideal maximum and the pretest scores. N-gain categories are determined based on Hake's criteria, namely high (> 0.70), moderate (0.30–0.70), and low (<0.30).

Based on Table 5 and Figure 2, students' mathematical problem-solving skills improved across all stages of Polya's method in the experimental class. The highest N-Gain score was observed in the "carrying out the plan" stage (0.95), falling into the high category, followed by the "understanding the problem" stage (0.77), which also fell into the high category. Both stages showed higher N-Gain scores than the control class at the corresponding stages (0.77 and 0.70), with the largest difference in the "carrying out the plan" stage. This indicates that Zep Quiz is most effective in encouraging students to apply problem-solving strategies systematically and accurately. Conversely, the N-Gain scores for the "devising a plan" stage (0.68) and the "looking back" stage (0.67) fell into the medium category, although they remained higher than those of the control class (0.64 and 0.57). The relatively low N-Gain scores in the "looking back" stage across both classes indicate that re-evaluating answers remains an area that requires greater instructional attention, whether using Zep Quiz or conventional teaching methods.

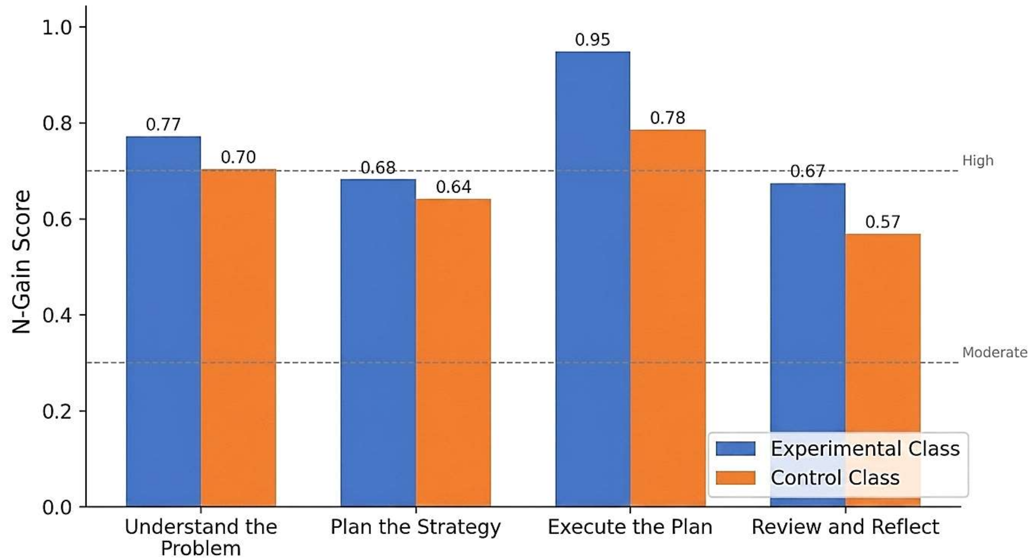

Figure 2. N-Gain scores per indicator (polya's stages): experimental class and control class. Note: The horizontal dashed line indicates the N-Gain criterion thresholds according to Hake (1998): scores e" 0.70 are categorized as high, and 0.30–0.70 as medium

Before the treatment, both classes had roughly equal initial abilities. The pretest results showed that the average score for the experimental class was 45.50 and for the control class was 45.18, with the highest score in both classes being 60; the lowest in the experimental class was 15, and the lowest in the control class was 20. This equality of initial abilities is an important requirement in quasi-experimental research because it ensures that differences in final results are attributable to the treatment rather than to differences in students' initial abilities (Abdullah et al., 2022). This condition is also consistent with the homogeneity test results, which showed that the N-gain data for both classes were homogeneous ($p = 0.408 > 0.05$), allowing a valid comparison between the two classes.

After the treatment, students' mathematical problem-solving abilities increased significantly in both classes. However, the increase experienced by the experimental class was much greater than that of the control class. In the experimental class, the average posttest score was 87.30, with a minimum of 70 and a maximum of 100. Meanwhile, the control class achieved an average posttest score of 81.93, with a minimum of 60 and a maximum of 100. This 5.37-point difference reflects the superiority of Zep Quiz-assisted learning in improving mastery of statistics material, especially problem-solving abilities. Zep Quiz: The following figure shows the application of Zep Quiz during the learning process. Figure 3 shows the results of group scores monitored based on student progress using the Zep Quiz application.

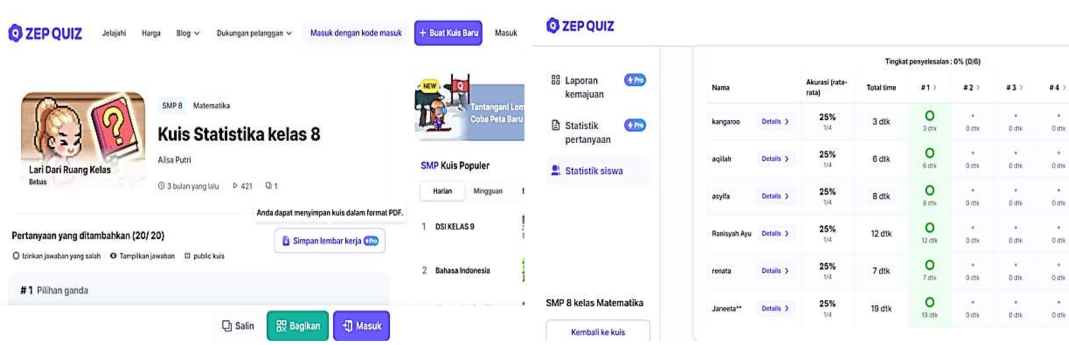

Figure 3. Example of practice scores monitored from student progress on the zep quiz application

Another indicator of effectiveness is the N-gain value, which represents the magnitude of the relative increase from the initial to the final condition after treatment. The experimental class obtained an N-gain score of 0.77, which falls in the high category, while the control class obtained an N-gain score of 0.66, which falls in the medium category. The N-gain category follows the Hake criteria: N-gain $\geq$ 0.70 is high, 0.30 d" N-gain $\leq$ 0.70 is medium, and N-gain < 0.30 is low. This difference in categories clearly shows that using Zep Quiz media produces superior learning effectiveness compared to conventional learning.

Another important aspect is the change in variance values in both classes. In the experimental class, the variance decreased significantly, from 126.46 on the pretest to 76.49 on the posttest. This decrease in variance indicates that after learning with Zep Quiz, students' problem-solving abilities became more evenly distributed, and the gap in scores between students narrowed significantly. This contrasted with the control class, where variance tended to remain stable, with a pretest variance of 161.95 and a posttest variance of 116.35. This indicates that the Zep Quiz media is not only effective in increasing the average student's ability but also in equalizing abilities among students. Thus, all students benefit more evenly from the learning process. This aligns with the view that interactive digital media can facilitate more inclusive and personalized learning (Yusnan, 2025).

Before conducting the hypothesis test, a series of prerequisite tests was conducted to ensure the data met the parametric statistical assumptions. The normality test was carried out using the Shapiro-Wilk test in SPSS. The normality test results showed that the post-test data in the experimental class had a significance value of 0.114 > 0.05, and the control class had a significance value of 0.156 > 0.05. Thus, H0 was accepted, and it was concluded that the post-test data in both classes were normally distributed. The normality test was conducted using the Shapiro-Wilk test in SPSS.

A homogeneity test was conducted using SPSS. The test results showed a significance value of 0.408 > 0.05, thus concluding that the N-gain data had homogeneous variance. Next, a linearity test was conducted to assess whether a linear relationship existed between the pretest data (independent variable) and the posttest data (dependent variable). The linearity test results showed a significance value of 0.246 > 0.05 for the experimental class and 0.171 > 0.05 for the control class, indicating a linear relationship between the pretest and posttest on problem-solving ability in both classes.

Because all prerequisite test assumptions were met, namely, the data were normally distributed, homogeneous, and linear, the hypothesis test was conducted using a parametric t-test with a significance level of á = 0.05. The hypothesis test showed a significance value (2-tailed) of 0.038 < 0.05, thus rejecting H0 and accepting H1. This means there is a significant difference in the mathematical problem-solving abilities of students who use the interactive learning media Zep Quiz and those who use conventional learning for statistics material in class VIII of SMPIT Al Qalam Bengkulu Selatan. This finding statistically proves that the Zep Quiz learning media is effective in improving students' mathematical problem-solving abilities. This finding reinforces the view of Liu et al. (2025), who emphasized that collaboration and discussion activities are very important for students in building an understanding of mathematical concepts. Student activities in using Zep Quiz media are shown in Figure 4.

To further enrich this discussion, I believe it is important to situate the current research findings within the broader body of literature on gamified quiz-based learning platforms, while also examining Zep Quiz in its own right. Several studies employing designs similar to the present

Figure 4. Student activity in completing the zep quiz media

one report generally consistent findings. Pellas (2024), using a quasi-experimental design with 12th-grade students in Greece, found that Kahoot! could significantly improve mathematics achievement compared to traditional instruction. Similarly, Jarrah et al. (2025), in their study involving 10th-grade students in the United Arab Emirates, reported that Kahoot! was highly effective for tasks requiring reasoning and problem-solving. This effect is attributed to its ability to maintain student engagement through competitive, real-time feedback. In the Indonesian context, Setiyani et al. (2020) employed a pretest-posttest design combined with N-gain analysis, an approach closely resembling the methodology of the present study, and found a significant difference in the improvement of mathematical problem-solving skills between students taught with Quizizz and those taught without it.

However, not all comparative evidence points to the same conclusion. For instance, Azzahra and Kowiyah (2022) compared Quizizz with Liveworksheets in a quasi-experimental study and found that students using Liveworksheets achieved higher post-test scores in mathematical problem-solving compared to those using Quizizz. This indicates that competitive, time-limited quiz formats do not necessarily outperform other digital tools. These differences suggest that a gamification platform's effectiveness for problem-solving skills depends not merely on the presence of game elements, but on how well its design features align with the cognitive demands of the assigned task.

Both Zep Quiz and Kahoot! emphasize real-time competition, with Zep Quiz further situating this competition within an explorable virtual environment. This may explain the convergent success observed for both platforms in the present study, as well as in the studies by Jarrah et al. (2024) and Pellas (2024). In contrast, Liveworksheet prioritizes structured, self-paced exercises. This approach appears better suited for consolidating problem-solving procedures before students are required to apply them under the pressure of competitive time constraints. This comparison suggests that the platforms' design characteristics, rather than gamification itself, moderate the extent of learning outcome improvements. Future research on Zep Quiz should investigate this factor in greater depth by comparing the platform with other digital tools (gamified or not) using the same student population.

From a gamification theory perspective, Zep Quiz's success in improving problem-solving skills in junior high school students can be explained through several psychological mechanisms. First, Zep Quiz's reward system, which includes scores and rankings, triggers dopamine release and makes students feel appreciated for their efforts, aligning with the characteristics of junior high school students, who are highly sensitive to rewards. Second, the game environment creates a safe space to fail, where not answering questions is not seen as embarrassing but as part of the learning process that requires retrying. Third, progress, which can be monitored through a dashboard, provides psychological satisfaction for students, motivating them to continue improving their abilities (Anggrayeni et al., 2025). This positive psychological state indirectly contributes to improved students' cognitive abilities, including mathematical problem-solving.

CONCLUSION

This study empirically demonstrates that using the interactive learning media Zep Quiz is significantly more effective at improving the mathematical problem-solving skills of eighth-grade students at SMPIT Al Qalam Bengkulu Selatan than conventional learning. This is evidenced by: (1) the N-gain of the experimental class of 0.77 (high category), which is greater than the N-gain of the control class of 0.66 (medium category); and (2) the results of the t-test with a sig. (2-tailed) value of $0.038 < 0.05$, indicating a statistically significant difference.

This study recommends that junior high school mathematics teachers explore using gamification platforms such as Zep Quiz, especially for materials that require high-level reasoning, such as statistics. Further research is recommended to examine the effectiveness of Zep Quiz across different mathematics materials and broader educational levels, and to integrate qualitative analysis to better understand students' learning experiences.

USE OF ARTIFICIAL INTELLIGENCE (AI)-ASSISTED TECHNOLOGY

The authors declare that no artificial intelligence (AI) was used in the preparation of this manuscript. All aspects of the research, including data collection, data analysis, interpretation, and manuscript preparation, were carried out entirely by the authors without the assistance of AI-based technology.

REFERENCES

Abdullah, K., Jannah, M., Aiman, U., Hasda, S., Fadilla, Z., Taqwin, Masita, Ardiawan, K. N., & Sari, M. E. (2022). *Metodologi penelitian kuantitatif* (N. Saputra, Ed.). Yayasan Penerbit Muhammad Zaini.

Amalina, I. K., & Vidákovich, T. (2023). Cognitive and socioeconomic factors that influence the mathematical problem-solving skills of students. *Heliyon*, *9*(9). https://doi.org/10.1016/j.heliyon.2023.e19539

Aprijal, A., Alfian, A., & Syarifudin, S. (2020). *Pengaruh minat belajar siswa terhadap hasil belajar siswa di Madrasah Ibtidaiyah Darussalam Sungai Salak Kecamatan Tempuling* [The influence of students' learning interest on student learning outcomes at Darussalam Sungai Salak Elementary School, Tempuling District]. *Mitra PGMI: Jurnal Kependidikan MI*, *6*(1), 76–91. https://doi.org/10.46963/mpgmi.v6i1.125

Anggrayeni, W., Hajar, I., & Hidayati, N. (2025). *Peningkatan motivasi berprestasi dengan gamifikasi pada mata pelajaran IPA di SMPN 6 Siak Hulu Kabupaten Kampar tahun ajaran 2024/2025* [Increasing achievement motivation with gamification in science subjects at SMPN 6 Siak Hulu, Kampar Regency, 2024/2025 academic year]. *Jurnal Pendidikan Sains dan Teknologi Terapan*, *2*(1), 26–33. https://jurnal.kopusindo.com/index.php/jpst/article/view/605

As Sa'adah, L. H., & Wulandari, E. (2022). Interest in learning mathematics for students of SMP Negeri 2 Tuntang Semarang

Regency towards learning during a pandemic. *AlphaMath/ : Journal of Mathematics Education*, *8*(2), 153. https://doi.org/10.30595/alphamath.v8i2.14321

Azzahra, M., & Kowiyah. (2022). The differences between using Quizizz interactive quiz and LiveWorksheet in math problem solving ability. *Jurnal Penelitian Dan Pengembangan Pendidikan*, *6*(2), 321–328. https://doi.org/10.23887/jppp.v6i2.48883

Bayaga, A. (2024). Enhancing M enhancing mathematics problem-solving skills in AI-driven environment: Integrated SEM-neural network approach. *Computers in Human Behavior Reports*, *16* (September), 100491. https://doi.org/10.1016/j.chbr.2024.100491

Bishara, S. (2016). Creativity in unique problem-solving in mathematics and its influence on motivation for learning. *Cogent Education*, *3*(1). https://doi.org/10.1080/2331186X.2016.1202604

Cãprioarã, D. (2015). Problem solving - purpose and means of learning mathematics in school. *Procedia - Social and Behavioral Sciences*, *191*, 1859–1864. https://doi.org/10.1016/j.sbspro.2015.04.332

Chaiarwut, S., Srikoon, S., Siritaratiwat, A., & Kwangmuang, P. (2025). Enhancing executive mathematics problem-solving through a constructivist digital learning platform: Design, development and evaluation. *Social Sciences and Humanities Open*, *11*(June 2024), 101338. https://doi.org/10.1016/j.ssaho.2025.101338

Chieng, B., Wen, S., & Sulaiman, N. A. (n.d.). Impact of ZEP QUIZ on vocabulary acquisition and engagement among Year 3 remedial ESL Learners: A quasi-experimental study. https://doi.org/10.47772/IJRISS.2026.10100320

Daryanes, F., Darmadi, D., Fikri, K., Sayuti, I., Rusandi, M. A., & Situmorang, D. D. B. (2023). The development of articulate storyline interactive learning media based on case methods to train student's problem-solving ability. *Heliyon*, *9*(4), e15082. https://doi.org/10.1016/j.heliyon.2023.e15082

Dewi, Y. M., Purnomo, E. A., & Sulistyaningsih, D. (2025). *Studi literatur review: Analisis kemampuan matematis siswa dalam menyelesaikan soal PISA ditinjau dari gaya kognitif field dependent dan field independent* [Literature review study: Analysis of students' mathematical abilities in solving PISA questions viewed from the perspective of field-dependent and field-independent cognitive styles]. *Jurnal Cendekia/ : Jurnal Pendidikan Matematika*, *9*(2), 788–799. https://doi.org/10.31004/cendekia.v9i2.3702

Fryer, L. K., Shum, A., King, R. B., & Lau, P. (2025). How does interest in a course interact with course learning? *Learning and Instruction*, *98*(March), 102106. https://doi.org/10.1016/j.learninstruc.2025.102106

Hariyati, A., & Purwanto, R. (2023). Implementing the contextual approach in mathematics learning to enhance students' problem solving abilities. *ASIAN: Indonesian Journal of Learning Development and Innovation*, *1*(1), 7–12. https://journal.institercom-edu.org/index.php/asian

Hasibuan, Z. (2021). The effect of learning models and learning interest on students' mathematics learning results on social arithmetic material for grade 7 Mtsn 3 Langkat. *Journal of Education and Practice*, *12*(33), 113–118. https://doi.org/10.7176/jep/12-33-13

Hilmi, I. N., & Jasiah. (2025). Revolusi penilaian pembelajaran: Implementasi Zep Quiz sebagai solusi digital interaktif di era pendidikan 4.0. *At-Tarbiyah: Jurnal Penelitian dan Pendidikan Agama Islam*, *3*(1), 608–614. https://journal.staittd.ac.id/index.php/at/article/view/649

Holmes, K., Mackenzie, E., Thompson, M., Reid O'Connor, B., & Anderson, J. (2025). Increasing student interest in secondary mathematics: student perspectives. *Mathematics Education Research Journal*. https://doi.org/10.1007/s13394-025-00553-0

Jarrah, A. M., Wardat, Y., Fidalgo, P., & Ali, N. (2025). Gamifying mathematics education through Kahoot: Fostering motivation and achievement in the classroom. *Research and Practice in Technology Enhanced Learning*, *20*, 1–27. https://doi.org/10.58459/rptel.2025.20010

Kopparla, M., Bicer, A., Vela, K., Lee, Y., Bevan, D., Kwon, H., Caldwell, C., Capraro, M. M., & Capraro, R. M. (2019). The effects of problem-posing intervention types on elementary students' problem-solving. *Educational Studies*, *45*(6), 708–725. https://doi.org/10.1080/03055698.2018.1509785

Larsson, M., Fredriksdotter, H., & Klang, N. (2024). Different types of talk in mixed-attainment problem-solving groups: Contributions to individual students' combinatorial thinking. *Journal of Mathematical Behavior*, *76*(January). https://doi.org/10.1016/j.jmathb.2024.101206

Liu, J., Sun, D., Sun, J., Wang, J., & Yu, P. L. H. (2025). Designing a generative AI enabled learning environment for mathematics word problem solving in primary schools: Learning performance, attitudes and interaction. *Computers and Education: Artificial Intelligence*, *9*(October 2024), 100438. https://doi.org/10.1016/j.caeai.2025.100438

Mägdefrau, J., Wecker, C., Engelmann, K., Michler, A., Baumgartner, I., & Böhm, M. (2025). Task-specific situational interest and appropriate learning strategy use in history lessons: A cross-classified multilevel analysis. *Learning and Instruction*, *98*(March). https://doi.org/10.1016/j.learninstruc.2025.102116

Masito, D., Tuhu, S., Morin, S., & Saputri, V. (2024). The relationship between indicators of mathematical problem-solving ability and performance on The PISA test: Systematic literature review. *Didaktika: Jurnal Kependidikan*, *13*(4), 5149–5164. https://doi.org/10.58230/27454312.1052

Mayoral-Rodríguez, S., Timoneda-Gallart, C., & Pérez-Álvarez, F. (2018). Effectiveness of experiential learning in improving cognitive planning and its impact on problem solving and mathematics performance. *Cultura y Educacion*, *30*(2), 308–337. https://doi.org/10.1080/11356405.2018.1457609

Montero-Izquierdo, A. I., Jeong, J. S., & González-Gómez, D. (2024). A future classroom lab with active and gamified STEAM proposal for mathematics and science disciplines: Analyzing the effects on pre-service teacher's affective domain. *Heliyon*, *10*(16). https://doi.org/10.1016/j.heliyon.2024.e35911

Munggaran, R. A.-A., Amalia, A. R., & Nurasiah, I. (2025). *Upaya peningkatan hasil belajar matematika melalui media tabel perkalian digital berbasis Wordwall* [Efforts to improve mathematics learning outcomes through digital multiplication table media based on Wordwall]. *Al-Irsyad Journal of Mathematics Education*, *4*(2), 451–458. https://doi.org/10.58917/

ijme.v4i2.368

Mustafa, S., Amaluddin, Nurhaeda, & Sari, V. (2023). Fun learning method in effecting the students ' interest in learning mathematics. Kreano*, Jurnal Matematika Kreatif-Inovatif, 14*(1), 12–27. https://doi.org/10.15294/kreano.v14i1.40588

Pathuddin, Rahmawati, S., Ikfal, Sukayasa, & M, B. (2025). The effect of time management and interest in learning mathematics: A case study of senior high school students in Palu Indonesia on students' learning achievement. *Heliyon*, *11*(3), e42048. https://doi.org/10.1016/j.heliyon.2025.e42048

Pellas, N. (2024). Effects of Kahoot! on K-12 Students' mathematics achievement and multi-Screen addiction. *Multimodal Technologies and Interaction*, *8*(9). https://doi.org/10.3390/mti8090081

Perche, L., Yennek, N., & Léger, L. (2025). Interest and mind wandering: How do individual and situational characteristics impact learning? *Learning and Motivation*, *92*(March). https://doi.org/10.1016/j.lmot.2025.102177

Ramadhani, A. P., Putri, H. E., Cahaya, D., Sari, K., Waffa, A., Shalimar, J., & Azhar, Z. J. (2025). *Efektivitas penggunaan Zep Quiz dalam meningkatkan literasi numerasi siswa SD* [The effectiveness of using Zep Quiz in improving elementary school students' numeracy literacy]. *Jurnal Kolaboratif Sains*, *8*(12), 8083–8095. https://doi.org/10.56338/jks.v8i12.9516

Riyanto N.A., & Amidi. (2024). *Studi literatur/ : Kemampuan pemecahan masalah matematis dalam model pembelajaran connecting, organizing, reflecting, extending (CORE)* [Literature study: Mathematical problem solving ability in the connecting, organizing, reflecting, extending (CORE) learning model]. *PRISMA, Prosiding Seminar Nasional Matematika*, *7*(1), 261–267.

Rocha, H., & Babo, A. (2024). Problem-solving and mathematical competence: A look to the relation during the study of linear programming. *Thinking Skills and Creativity*, *51*(June 2023), 101461. https://doi.org/10.1016/j.tsc.2023.101461

Salamah, A. S., Gianova, A., Nabila, D. P., Utami, R. S. P., Sabili, T. H., & Putri, H. E. (2025). *Pengaruh penggunaan media evaluasi digital Zep Quiz terhadap hasil belajar matematika materi bilangan cacah siswa kelas V SDN 6 Nagri Kaler Purwakarta* [The effect of using the digital evaluation media Zep Quiz on the mathematics learning outcomes of fifth grade students of SDN 6 Nagri Kaler Purwakarta in the material of natural numbers.]. *Jurnal Review Pendidikan Dan Pengajaran*, *8*(4), 8814–8822. https://doi.org/10.31004/jrpp.v8i4.53961

Sari, C. N., Sutini, S., & Prasetyo, A. (2025). Pengaruh Zep Quiz sebagai digital learning media terhadap hasil belajar matematika siswa Upt SMPN 2 Gresik. *De Fermat/ : Jurnal Pendidikan Matematika*, *8*(2), 464–472. https://doi.org/10.36277/defermat.v8i2.2327

Sausan, T., & Wibowo, M. U. (2024). *Analisis kemampuan literasi matematis peserta didik sekolah menengah dalam menyelesaikan soal PISA konten quantity ditinjau dari math anxiety* [Analysis of the mathematical literacy abilities of secondary school students in solving PISA questions on quantity content viewed from the perspective of math anxiety]. *Jurnal Agama dan Sosial Humaniora*, *1*(1), 18–33. https://doi.org/10.15575/jash.v1i1.713

Schreiber, I., & Ashkenazi, H. (2024). The impact of various teaching methods on the knowledge of students of different genders: The case of mathematics word problems abstract. *Heliyon*, *10*(16), e35610. https://doi.org/10.1016/j.heliyon.2024.e35610

Setiyani, S., Fitriyani, N., & Sagita, L. (2020). Improving student's mathematical problem solving skills through Quizizz. *JRAMathEdu (Journal of Research and Advances in Mathematics Education)*, *5*(3), 276–288. https://doi.org/10.23917/jramathedu.v5i3.10696

Sriwahyuni, K., & Maryati, I. (2022). *Kemampuan pemecahan masalah matematis siswa pada materi statistika* [Students' mathematical problem-solving abilities in statistics material]. *Plusminus: Jurnal Pendidikan Matematika*, *2(2),335–334*. https://doi.org/10.31980/plusminus.v2i2.1109

Turner, E., Gutiérrez, R. J., & Sutton, T. (2011). Student participation in collective problem solving in an after-school mathematics club: Connections to learning and identity. *Canadian Journal of Science, Mathematics and Technology Education*, *11*(3), 226–246. https://doi.org/10.1080/14926156.2011.595884

Wilkie, K. J. (2024). Creative thinking for learning algebra: Year 10 students' problem solving and problem posing with quadratic figural patterns. *Thinking Skills and Creativity*, *52*(May), 101550. https://doi.org/10.1016/j.tsc.2024.101550

Williams, K., & Williams, H. (2021). Mathematics problem-solving homework as a conduit for parental involvement in learning. Evaluation of a pilot study. *Educational Review*, *73*(2), 209–228. https://doi.org/10.1080/00131911.2019.1566210

Yeribatuah, P., & Arthur, Y. D. (2023). Learners' attitude towards mathematics, technology on their mathematics interest: The Mediating Role of Learners' Mathematics Perception. *Mathematics Education Journal*, *17*(3), 361–378. https://doi.org/10.22342/jpm.17.3.22103.361-378

Yuliati, I. (2021). *Kemampuan pemecahan masalah matematika ditinjau dari minat belajar peserta didik* [mathematical problem-solving ability reviewed from students' learning interests]. *Jurnal Cendekia/ : Jurnal Pendidikan Matematika*, *5*(2), 1159–1168. https://doi.org/10.31004/cendekia.v5i2.547

Yusnan, M. (2025). *Media pembelajaran interaktif*. Eureka Media Aksara.