Kognitif
Jurnal Riset HOTS Pendidikan Matematika
Volume - 6 No - 3 2026 Page 1256 – 1273
ISSN 2776-9704 P-ISSN 2776-9984
https://doi.org/10.51574/kognitif.v6i3.5637

Development of a Problem-Solving-Based Mathematics Educational Game to Support High School Students' Mathematical Reasoning Ability

Nisa Aghnia, Novaliyosi, Maman Fathurrohman

How to cite: Aghnia, N., Novaliyosi, N., & Fathurrohman, M. (2026). Development of a Problem-Solving-Based Mathematics Educational Game to Support High School Students' Mathematical Reasoning Ability. *Kognitif: Jurnal Riset HOTS Pendidikan Matematika*, *6*(3), 1256–1273. https://doi.org/10.51574/kognitif.v6i3.5637

To link to this artcle: https://doi.org/10.51574/kognitif.v6i3.5637

Opened Access Article

Published Online on 24 August 2026

Submit your paper to this journal

Kognitif: Jurnal Riset HOTS Pendidikan Matematika
https://doi.org/10.51574/kognitif.v6i3.5637 Volume 6, No 3, July - September 2026, pp. 1256 –1273

Development of a Problem-Solving-Based Mathematics Educational Game to Support High School Students' Mathematical Reasoning Ability

Nisa Aghnia[1*], Novaliyosi, Maman Fathurrohman

[1]Department of Mathematics Education, Faculty of Teacher Training and Education, Universitas Sultan Ageng Tirtayasa

Article Info

Article history:

Received Aug 08, 2026
Accepted Aug 19, 2026
Published Online Aug 24, 2026

Keywords:

ADDIE
Educational Game
Mathematical Reasoning Ability
Problem Solving
Scratch

ABSTRACT

This study was conducted to provide an interactive problem-solving-based learning medium that supports the development of high school students' mathematical reasoning ability. This study aimed to develop a mathematics educational game using Scratch on statistics material for tenth-grade high school students and to determine the validity, practicality, and effectiveness of the product. This study employed a *Research and Development* (R&D) method using the ADDIE development model. The participants consisted of three material experts, three media experts, one mathematics teacher, 35 students in the limited trial, and 32 students in the main field trial at SMA Negeri 5 Cilegon. Data were collected through validation sheets, response questionnaires, and a mathematical reasoning ability test, and were analyzed using percentages and Normalized Gain (N-Gain). The results showed that the developed game obtained a validity score of 90% from material experts and 89% from media experts, both categorized as very valid. The game also met the practicality criteria based on the teacher's and students' responses. The use of the game resulted in an N-Gain score of 0.567, indicating a moderate increase in students' mathematical reasoning ability. Thus, the problem-solving-based educational game was considered valid and practical and showed potential to support students' mathematical reasoning ability in statistics learning.

Corresponding Author:

Nisa Aghnia,
Department of Mathematics Education,
Faculty of Teacher Training and Education,
Universitas Sultan Ageng Tirtayasa,
No. 25, Ciwaru Raya Street, Cipare, Serang Sub-district, Serang City, Banten 42117, Indonesia
nisaaghnia705@gmail.com

Introduction

Mathematical reasoning ability is one of the essential abilities in mathematics learning because it helps students construct arguments, provide justifications, and draw logical conclusions (National Council of Teachers of Mathematics, 2000). Mathematical reasoning ability is also associated with students' ability to present mathematical statements, construct proofs or solutions, and draw logical conclusions (Prastika et al., 2021). Mathematical reasoning plays an important role in the problem-solving process because students need to connect information, construct arguments, and draw conclusions based on logical thinking processes (Agustin et al., 2021). The development of these abilities is in line with the implementation of the Merdeka Curriculum, which emphasizes reasoning and problem-solving abilities as part of Higher Order Thinking Skills (HOTS) (Kemendikbudristek, 2022).

The mathematical achievement of Indonesian students remains a concern. The results of the 2022 Programme for International Student Assessment (PISA) showed that Indonesian students' mathematical literacy was still below the average of OECD member countries. Mathematical literacy in PISA encompasses broader abilities than mathematical reasoning; therefore, these results cannot be used as direct evidence of low mathematical reasoning ability. Research findings regarding students' difficulties in dealing with problems that require reasoning, information interpretation, and contextual problem solving also indicate that developing mathematical thinking remains a challenge (Mariza et al., 2022; Anggraini et al., 2022).

These problems were more specifically identified in mathematics learning at SMA Negeri 5 Cilegon. Initial observations showed that learning was still dominated by the use of textbooks, student worksheets, and presentation media, while the use of interactive learning media had not been optimized. Interviews with teachers and students indicated that students still experienced difficulties in determining solution strategies, providing reasons for their answers, and connecting mathematical concepts when solving problems. The pretest results of 32 students showed an average score of 51.56, with the lowest achievement found in the indicator of making conjectures at 39.06% and constructing proofs or providing reasons for the correctness of solutions at 45.31%. These conditions indicate the need for learning that not only provides materials and final answers but also facilitates students' thinking processes in solving problems.

Problem-based learning provides students with opportunities to engage in mathematical problem-solving processes that can support the development of mathematical reasoning ability (Fitriyah et al., 2022; Suryanti et al., 2022). These processes are consistent with Polya's problem-solving steps, namely understanding the problem, devising a plan, carrying out the plan, and looking back at the solution obtained (Putra, Novaliyosi, et al., 2025). This structure can be used to operationalize mathematical reasoning processes because each stage of problem solving can be directed toward activities involving identifying information, making conjectures, determining strategies, providing justifications, checking the validity of solutions, and drawing conclusions.

Interactive learning media are needed so that problem-solving processes can be realized through learning activities that allow students to interact directly with problems. Scratch can be used to develop mathematics learning media that integrate text, images, animation, sound, and feedback within a single learning environment. The use of Scratch in mathematics learning has shown positive results. Indahwati et al. (2024) reported that Scratch-based learning media received positive responses from users. Indah Widiyanti et al. (2025) showed that Scratch-assisted student worksheets based on Problem-Based Learning met the criteria of validity, practicality, and sufficient effectiveness in improving mathematical problem-solving skills. Faradilla et al., (2023) reported that Scratch-based multimedia learning improved students' mathematical reasoning ability with an N-Gain in the moderate category. Salamah et al., (2024) reported very good responses to Android-based learning media developed using Scratch. Maulidina Rahadi et al., (2024) found that Scratch-assisted learning media met the criteria of being highly valid, practical, and effective and increased student engagement. Ardiyanti et al. (2025) showed that Scratch-based interactive learning media were highly feasible for mathematics learning.

Kognitif: Jurnal Riset HOTS Pendidikan Matematika
https://doi.org/10.51574/kognitif.v6i3.5637 Volume 6, No 3, July - September 2026, pp. 1256 –1273

These studies demonstrate the potential of Scratch to support mathematics learning, including the development of mathematical reasoning ability. However, the issue that remains to be investigated is not merely the use of Scratch for particular materials or educational levels, but how activities within Scratch-based media can be designed so that systematic problem-solving processes directly facilitate indicators of mathematical reasoning ability. For example, Faradilla et al., (2023), demonstrated an improvement in mathematical reasoning ability through Scratch-based multimedia, but the study did not specifically explain how each reasoning indicator was integrated into the problem-solving stages within media activities. Other studies using Scratch have focused more on validity, practicality, learning outcomes, problem-solving skills, user responses, or student engagement (Indah Widiyanti et al., 2025; Salamah et al., 2024; Maulidina Rahadi et al., 2024; Ardiyanti et al., 2025). This gap indicates the need for media that not only uses Scratch as a platform for presenting learning content but also organizes game activities based on an explicit relationship between mathematical reasoning indicators and problem-solving stages.

This study addresses this gap through the development of a problem-solving-based mathematics educational game using Scratch that integrates five indicators of mathematical reasoning ability with the four problem-solving stages proposed by Polya. This integration is operationalized through a five-level game structure, in which each level focuses on one mathematical reasoning indicator and requires students to solve problems through the stages of understanding the problem, devising a plan, carrying out the plan, and looking back at the solution. The game activities do not stop at selecting an answer but guide students through thinking processes corresponding to the reasoning indicator addressed at each level. Scaffolding based on Vygotsky's social constructivist theory is implemented through hints, feedback, and teacher assistance to support students when they experience difficulties in completing the problem-solving stages. Scaffolding provides gradual learning support according to students' needs, thereby helping them construct understanding and solve problems more independently (Saputra et al., 2024). In the game design, such support functions as assistance during the problem-solving process rather than as a replacement for students' thinking processes. The integration of reasoning indicators, Polya's problem-solving stages, and scaffolding constitutes a characteristic that distinguishes the developed game from Scratch-based media that primarily function as platforms for presenting materials, exercises, or educational games. Therefore, this study aimed to develop a problem-solving-based mathematics educational game using Scratch on statistics material for tenth-grade high school students and to determine the validity, practicality, and effectiveness of the product. The developed product is expected to provide an alternative mathematics learning medium that offers a structured problem-solving experience and supports the development of students' mathematical reasoning ability.

Method

Research Design and Development Procedure

This study employed a Research and Development (R&D) method to develop a problem-solving-based mathematics educational game using Scratch on statistics material for tenth-grade high school students. The development model used was ADDIE (Analysis, Design, Development, Implementation, and Evaluation), developed by Branch. The ADDIE model was selected because it provides a systematic development process from needs analysis to product evaluation.

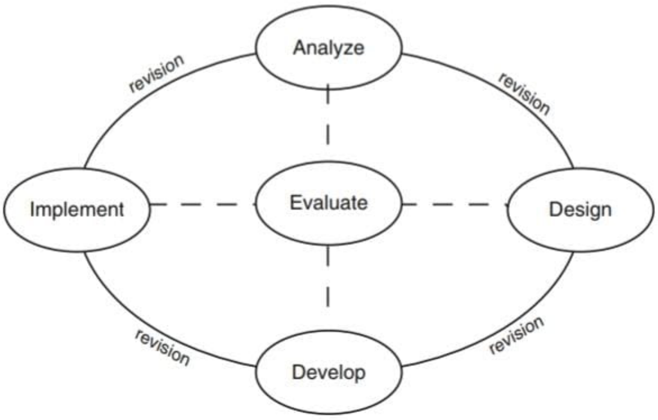

Figure 1. ADDIE Development Model
Source: Branch (2009)

The analysis stage involved needs analysis, analysis of student characteristics, curriculum analysis, observation, interviews, and a literature review. The design stage included the preparation of the storyboard, interface design, statistics materials, and research instruments. The development stage involved developing the game using Scratch, followed by validation by material and media experts and revision based on the validators' feedback. The implementation stage was conducted through a limited trial and a main field trial to obtain practicality data and measure changes in students' mathematical reasoning scores after using the game. The evaluation stage was conducted formatively at each development stage as a basis for product improvement.

Participants

The study was conducted at SMA Negeri 5 Cilegon during the second semester of the 2025/2026 academic year. Participants were selected using purposive sampling based on the suitability of their characteristics with the objectives and stages of the study. The selection of trial classes was based on the initial observations and recommendations of the mathematics teacher. The participants consisted of three material experts, three media experts, one mathematics teacher, students involved in the instrument trial, students involved in the limited trial, and students involved in the main field trial. The material experts consisted of two Mathematics Education lecturers and one mathematics teacher. The media experts consisted of one Information and Communication Technology teacher, one Mathematics Education lecturer, and one Science Education lecturer. The material experts assessed the suitability of the content, conceptual accuracy, presentation, language, and alignment of the material with the learning objectives and mathematical reasoning indicators. The media experts assessed the aspects of appearance, interactivity, navigation, programming, ease of use, and overall media functionality. The instrument trial involved 30 students from class XI-3. The limited trial involved 35 students from class X-4 and was conducted to obtain initial information regarding the practicality of the game and to identify product deficiencies before refinement. The main field trial involved 32 students from class X-6 and was conducted after the product had been revised based on the results of the limited trial.

Kognitif: Jurnal Riset HOTS Pendidikan Matematika
https://doi.org/10.51574/kognitif.v6i3.5637 Volume 6, No 3, July - September 2026, pp. 1256 –1273

Research Instruments

The research instruments consisted of non-test and test instruments. The non-test instruments included material expert validation sheets, media expert validation sheets, teacher response questionnaires, and student response questionnaires. The test instrument consisted of essay questions measuring mathematical reasoning ability and was used to obtain data on students' abilities before and after using the game. The mathematical reasoning ability test consisted of five essay items, with each item representing one indicator of mathematical reasoning ability. The instrument was tested on 30 students from class XI-3 before being used in the main study. Instrument quality was analyzed in terms of item validity, reliability, difficulty level, and discrimination power.

Test Instrument Blueprint

Table 1. Mathematical Reasoning Ability Test Indicators and Test Items

Mathematical Reasoning Indicator	Sample Item
Making conjectures Students identify the given and required information and formulate an initial conjecture based on the data presented	A sports coach records the average running time of athletes in a group. Several new running-time data points are then added to the record. Questions: a. Without performing any calculations, how do you think the average running time might change after the new data are added? Write down the possible changes that may occur and briefly explain your reasoning. b. From the possibilities you have identified, explain how you could determine which possibility is applicable. c. Choose two different possibilities and provide an example for each. Explain how each example affects the average running time. Based on your findings, explain the general conditions that determine how the average running time changes. An athlete states: "Adding new data will definitely make the average running time higher." Does this statement agree with your conclusion? Explain your answer
Performing mathematical manipulations Students determine a solution strategy and appropriately apply statistical concepts, rules, or procedures to solve the problem.	A teacher knows that the mean score of a mathematics test for 8 students is 75. The seven known student scores are 60, 65, 70, 75, 80, 85, and 90. One student's score is unknown. Questions : a. Identify the information that is known and what needs to be determined. b. Explain the relationship between the mean, the total score, and the number of data points that can be used to solve the problem. c. Suppose the unknown student's score is represented by . Form an equation that represents the given situation, then determine the value of and explain the steps used to solve it. Explain how changes in the value of can affect the mean. Determine the conditions under which the mean increases, decreases, or remains unchanged.

<table>
<tr><th>Mathematical Reasoning Indicator</th><th>Sample Item</th></tr>
<tr><td>Constructing proofs or providing reasons for the correctness of a solution

Students complete the calculations and provide logical reasons so that the obtained solution can be justified</td><td>A teacher states that the mean test score in the class is 80. There are 25 students in the class. Of these, 10 students scored 90, while the remaining 15 students obtained the same score. One student argues that the scores of the 15 students are less than 80.
Questions :
a. Identify the information that is known and what needs to be proven.
b. Explain the method or strategy that can be used to test the statement.
c. Determine whether the statement is correct by constructing a mathematical proof.
Explain why the proof you provide is sufficient to support your conclusion.</td></tr>
<tr><td>Examining the validity of an argument

Students evaluate the correctness of a statement or argument based on mathematical concepts and principles</td><td>A teacher records the test scores of 30 students as follows:
• 6 students scored 60.
• 8 students scored 70.
• 10 students scored 80.
• 6 students scored 90.
One student states that the class mean score is 77, which means that the mean is higher than 75. However, another student argues that the class mean should be close to 75 because most of the scores are in the middle range.
Determine which argument is correct and explain the mathematical reasoning that supports your answer.
Questions :
a. Identify the information that is known and what needs to be determined.
b. Explain the strategy that can be used to examine the validity of both arguments.
c. Perform the calculation and determine which argument is correct.
Explain why the argument you selected is appropriate and whether it is consistent with the given data.</td></tr>
<tr><td>Drawing logical conclusions

Students formulate conclusions consistent with the solution based on relevant reasons and evidence</td><td>Consider the following mathematics test score data for 50 students.
The minimum passing criterion (KKM) is 75.
<table><tr><th>Score Interval</th><th>Number of Students</th></tr><tr><td>50-59</td><td>5</td></tr><tr><td>60-69</td><td>10</td></tr><tr><td>70-79</td><td>20</td></tr><tr><td>80-89</td><td>10</td></tr><tr><td>90-99</td><td>5</td></tr></table>
Based on the data in the table, can it be concluded that learning in the class has been effective because most students have achieved the KKM? Explain your answer by considering that the data are presented in class intervals.
Questions :
a. Identify the information provided in the table and what needs to be determined regarding students' learning mastery.
b. Explain how to determine whether most students have achieved the KKM.</td></tr>
</table>

Kognitif: Jurnal Riset HOTS Pendidikan Matematika
https://doi.org/10.51574/kognitif.v6i3.5637 Volume 6, No 3, July - September 2026, pp. 1256 –1273

Mathematical Reasoning Indicator	Sample Item
	c. Use the data in the table to determine the number of students who have achieved and have not achieved the KKM, and compare the two groups. Based on the results, draw a conclusion about whether the statement that learning has been effective is justified, and explain the reasons supporting your conclusion.

Item Validity

Item validity was analyzed using the Pearson Product-Moment correlation. The analysis involved 30 students with a significance level of 5%. An item was considered valid when the calculated correlation coefficient exceeded the critical value. The interpretation of the correlation coefficient referred to Efendi et al. (2024).

Table 2. Results of the Test Item Validity Analysis

Item	r-calculated	r-critical	Interpretation	Decision
1	0,762		High	Valid
2	0,872		Very High	Valid
3	0,833	0,361	Very High	Valid
4	0,798		High	Valid
5	0,685		High	Valid

Based on Table 2. Results of the Test Item Validity Analysis, all five items met the validity criterion and were therefore considered valid for measuring students' mathematical reasoning ability.

Reliability

The reliability of the instrument was analyzed using Cronbach's Alpha. The interpretation of the reliability coefficient referred to (Nursalam et al., 2017).

Table 3. Results of the Instrument Reliability Test

Cronbach's Alpha	Interpretation	Decision
0,846	Very High	Reliable

The Cronbach's Alpha coefficient of 0.846 was categorized as very high. Therefore, the mathematical reasoning ability test instrument was considered reliable.

Difficulty Level

Item difficulty was analyzed to determine the difficulty level of each test item. The difficulty criteria referred to (Son, 2019).

Table 4. Results of the Item Difficulty Analysis

Item	Difficulty Level	Interpretation
1	0,433	Moderate
2	0,6161	Moderate
3	0,658	Moderate
4	0,733	Moderate
5	0,725	Moderate

Based on Error! Reference source not found., all test items were categorized as having a moderate difficulty level. Therefore, the five items met the difficulty criteria for measuring students' mathematical reasoning ability.

Discrimination Power

Item discrimination power was analyzed to determine the ability of each test item to distinguish students based on the measured ability. The discrimination criteria referred to (Son, 2019).

Table 5. Results of the Item Discrimination Power Analysis

Item	Discrimination Power	Interpretation
1	0,233	Fair
2	0,266	Fair
3	0,283	Fair
4	0,3	Fair
5	0,216	Fair

Based on Table 5, all test items had fair discrimination power. Therefore, all five items met the discrimination criteria used in this study.

Research Procedure and Data Collection

The research procedure followed the five stages of the ADDIE model. The analysis stage involved needs analysis, analysis of student characteristics, curriculum analysis, observation, interviews, and a literature review. The design stage included the preparation of the storyboard, interface design, statistics materials, and research instruments. The development stage involved developing the game using Scratch, validating the product through material and media experts, and revising the product based on their feedback. The implementation stage consisted of a limited trial and a main field trial. The evaluation stage was conducted formatively at each development stage to support product refinement. The limited trial involved 35 students from class X-4, whereas the main field trial involved 32 students from class X-6 at SMA Negeri 5 Cilegon. The limited trial was conducted to obtain initial responses regarding the practicality of the game and to identify product deficiencies before the main field trial. The main field trial was conducted to obtain practicality data and measure changes in students' mathematical reasoning scores after using the game.

The game was used in one learning session on statistics, with the mathematics teacher serving as a facilitator during the trial. Each game level had a time limit of 593 seconds (9 minutes 53 seconds). Each level was designed around Polya's four problem-solving stages: understanding the problem, devising a plan, carrying out the plan, and looking back. Students solved the problems in each level sequentially according to these stages. Students' mathematical reasoning ability was measured using a one-group pretest-posttest design. Students completed the pretest before using the game, used the game during the trial, and then completed the posttest. The pretest and posttest used two equivalent test sets based on the same five mathematical reasoning indicators but with different contexts or data.

Kognitif: Jurnal Riset HOTS Pendidikan Matematika
https://doi.org/10.51574/kognitif.v6i3.5637 Volume 6, No 3, July - September 2026, pp. 1256 –1273

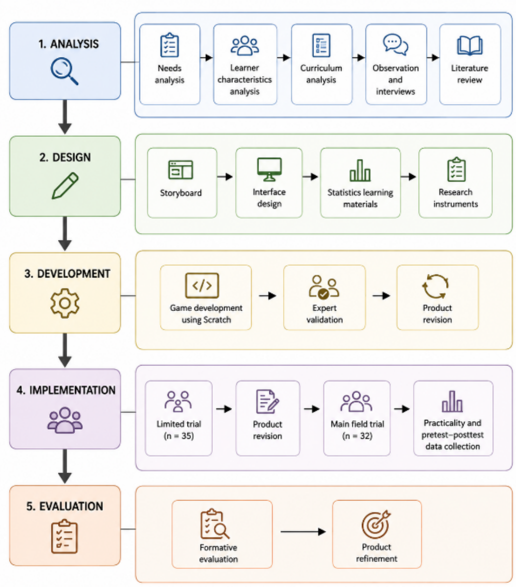

Figure 2. Research Procedure Flow

Data Analysis

The research data were analyzed qualitatively and quantitatively. Qualitative analysis was conducted on observations, interviews, documentation, and suggestions or feedback from material experts, media experts, the teacher, and students. The results were used as a basis for revising and refining the product at each development stage. Quantitative analysis was used to determine product validity and practicality and to analyze changes in students' mathematical reasoning scores after using the game. Data from expert validation and teacher and student response questionnaires were analyzed using a five-point Likert scale and converted into percentages.

Product Practicality

Product practicality was measured through three sources: teacher response, the limited trial, and the main field trial. The teacher response was obtained from one mathematics teacher, while the limited trial involved 35 students from class X-4 and the main field trial involved 32 students from class X-6. Practicality data were collected using response questionnaires with a five-point Likert scale. The practicality percentage was calculated based on the obtained score relative to the maximum possible score.

$$\text{Practicality} = \frac{\text{Total Score Obtained}}{\text{Maximum Possible Score}} \times 100\%$$

Table 6. Product Practicality Criteria

Score	Category
$81{,}25\% < \text{Score} \leq 100\%$	Very Practical
$62{,}50\% < \text{Score} \leq 81{,}25\%$	Practical
$43{,}75\% < \text{SScore} \leq 62{,}50\%$	Moderately Practical
$25\% < \text{Score} \leq 43{,}75\%$	Less Practical

Source: (Rahmania et al., 2023)

Improvement in Mathematical Reasoning Ability

The improvement in mathematical reasoning ability was analyzed based on the pretest and posttest scores from the main field trial involving 32 students. The study employed a one-group pretest-posttest design without a comparison group. The magnitude of improvement was analyzed using Normalized Gain (N-Gain).

$$g = \frac{S_{post} - S_{pre}}{S_{maks} - S_{pre}}$$

Notes:
g = N-Gain score
S_{pre} = Pre-test score
S_{post} = Post-test score
S_{maks} = Maximum score

Table 7. N-Gain Interpretation Criteria

Score Range	Criteria
$N - gain > 0,7$	High
$0,3 \leq N - gain \leq 0,7$	Moderate
$N - gain < 0,3$	Low

Source: (Rahmania et al., 2023)

The N-Gain score was used to describe the magnitude of improvement in students' mathematical reasoning scores after using the game. Because the study did not employ a comparison group, the N-Gain results were not used to establish a causal relationship between the use of the game and the improvement in mathematical reasoning ability.

Research Findings

This study produced a problem-solving-based mathematics educational game using Scratch on statistics material for tenth-grade high school students. The research results include product development, validity, practicality, and improvement in students' mathematical reasoning scores after using the product.

Product Development

The mathematics educational game was developed in the form of an adventure game that integrates statistics material with five indicators of mathematical reasoning ability through problem-solving activities at each game level. Each level was designed based on Polya's four stages of problem solving, namely understanding the problem, devising a plan, carrying out the plan, and looking back at the solution. The game consists of four main menus: Game, Instructions, Materials, and Developer Profile. The Game menu is used to start the game, the Instructions menu provides guiandce on how to use the game, the Materials menu contains a summary of statistics material, and the Developer Profile menu provides information about the media developer.

Kognitif: Jurnal Riset HOTS Pendidikan Matematika
https://doi.org/10.51574/kognitif.v6i3.5637 Volume 6, No 3, July - September 2026, pp. 1256 –1273

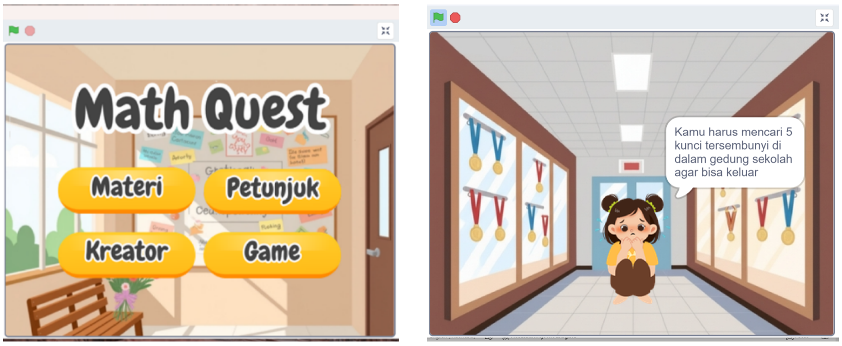

Figure 3. Main Menu Interface

The developed product was subsequently validated by material and media experts before being implemented.

Validity Test Results

The validity test was conducted by three material experts and three media experts. The assessment results are presented in Table 8 and Table 9.

Table 8. Results of Material Expert Validation

Validator	Percentage	Category
Material Expert 1	80%	Valid
Material Expert 2	93%	Very Valid
Material Expert 3	97%	Very Valid
Average	**90%**	**Very Valid**

Table 9. Results of Media Expert Validation

Validator	Percentage	Category
Material Expert 1	83%	Very Valid
Material Expert 2	88%	Very Valid
Material Expert 3	97%	Very Valid
Average	**89%**	**Very Valid**

The material experts' assessments ranged from 80% to 97%, with an average of 90%, categorized as very valid. The media experts' assessments ranged from 83% to 97%, with an average of 89%, also categorized as very valid. These results indicate that the product met the validity criteria in terms of both content and media aspects. The validators' feedback was used as a basis for refining the product before implementation. The revisions included aligning the test items with the mathematical reasoning indicators, adding contextual problems, adding illustrations and usage instructions, and improving the text presentation.

Practicality Test Results

The practicality of the game was assessed through the mathematics teacher's response, a limited trial, and a main field trial. The limited trial involved 35 students, while the main field trial involved 32 students. The assessment results are presented in Table 10.

Table 10. Results of the Practicality Test of the Educational Game

Respondent	Number	Percentage	Category
Mathematics Teacher	1	74%	Practical
Limited Trial	35	74%	Practical
Main Field Trial	32	68%	Practical

The teacher response obtained a percentage of 74%, while the limited trial obtained 74% and the main field trial obtained 68%. Although the percentage decreased in the main field trial, the result remained within the practical category based on the criteria used. This difference indicates variation in students' responses across the two trial stages, but it did not cause the product to fall below the established practicality threshold.

Implementation of Problem Solving and Scaffolding

The game implementation integrated five mathematical reasoning indicators into five game levels. Each level was completed through Polya's four stages, so students were not only required to provide final answers but were also guided to understand the problem, devise a plan, carry out the plan, and look back at the solution. The hint and feedback features were used as forms of scaffolding to provide assistance when students experienced difficulties during the problem-solving process. The teacher acted as a facilitator during the use of the game.

Figure 4. Implementation of the Game in Learning

Improvement in Mathematical Reasoning Ability

The improvement in students' mathematical reasoning scores was analyzed based on the pretest and posttest results from the main field trial involving 32 students. The results are presented in Table 11.

Table 11. Pretest and Posttest Results

Data	Score
Mean Pretest	51,56
Mean Posttest	79,19

The mean score increased from 51.56 on the pretest to 79.19 on the posttest, representing an increase of 27.63 points after the use of the game. The improvement was subsequently analyzed using Normalized Gain (N-Gain)

Table 12. N-Gain Analysis Results

N-Gain	Percentage	Category
0,567	56,69%	Moderate

The N-Gain value of 0.567 indicates that the improvement in students' mathematical reasoning scores after using the game was in the moderate category. These results indicate an improvement in scores in the main field trial group. However, the results cannot be used to establish a causal relationship between the use of the game and the improvement in mathematical reasoning ability because the study did not include a comparison group.

Kognitif: Jurnal Riset HOTS Pendidikan Matematika
https://doi.org/10.51574/kognitif.v6i3.5637 Volume 6, No 3, July - September 2026, pp. 1256 –1273

Discussion

Validity of the Educational Game

The validation results indicate that the problem-solving-based mathematics educational game obtained a very valid category. These results indicate that the product met the feasibility aspects of both the content and media for use in the trial. Product validity was not limited to the appearance and functionality of the media but also included the suitability of the statistics material, problem-solving activities, and mathematical reasoning indicators integrated into the game. The material expert validation instrument also included aspects related to students' ability to make conjectures, provide mathematical reasoning, draw conclusions, evaluate the correctness of solutions, and use various strategies. The validation process provided feedback that was used to refine the product before implementation. One important revision involved aligning the mathematical reasoning indicators presented in the game with the indicators used in the study, namely making conjectures, performing mathematical manipulations, constructing proofs or providing reasons for the correctness of solutions, examining the validity of arguments, and drawing logical conclusions. These findings indicate that validation in this study served not only to obtain feasibility scores but also to ensure that the development objectives were reflected in the product design. The results are consistent with Syarifah et al. (2025), Novaliyosi et al. (2022), and Fortuna et al. (2021), who demonstrated the importance of expert validation in the development of instructional media or learning resources. The present study provides a more specific contribution because the validation was directed not only toward the feasibility of Scratch-based media but also toward the alignment of the game design with the problem-solving process and mathematical reasoning ability.

Practicality of the Educational Game

The practicality test results indicate that the game was categorized as practical based on the teacher's response, limited trial, and main field trial. These results indicate that the game could be used by students to engage in game activities and solve the problems provided. The practicality of the product was supported by the availability of instructions, materials, and a systematically organized game flow, enabling students to follow the activities without requiring continuous explanations from the teacher. Although all stages were categorized as practical, the practicality percentage decreased from 74% in the limited trial to 68% in the main field trial. This difference is important to consider because it indicates that user responses were not entirely consistent when the product was used with a larger group. The implementation results indicate that some aspects received lower responses, which may be related to students' adaptation to the game mechanics, navigation, and problem-solving stages. Therefore, the practical category does not mean that the product was free from usability constraints; rather, it indicates that the identified constraints did not cause the product to fall below the established practicality threshold. These findings are consistent with Novaliyosi et al. (2022), Putra et al. (2025), and Monalisa et al. (2025), who demonstrated that technology-based learning media or educational games can meet practicality criteria. The distinction of the present study lies in the use of the game not merely as a medium for presenting learning materials but as a problem-solving environment that guides students through Polya's stages. Thus, practicality in this study relates to students' ability to use the media while simultaneously following the problem-solving process that has been designed.

Improvement in Mathematical Reasoning Ability

The main field trial results showed that the mean mathematical reasoning score increased from 51.56 on the pretest to 79.19 on the posttest, with an N-Gain of 0.567, which was categorized as moderate. These results indicate an increase in mathematical reasoning scores after the use of the game within the context of this study. However, the improvement cannot be interpreted as evidence that the game causally caused the improvement in mathematical reasoning because the study did not include a comparison group. The students' initial performance across the five mathematical reasoning indicators varied. The indicator of making conjectures obtained a score of 39.06%, constructing proofs or providing reasons for the correctness of solutions obtained 45.31%, performing mathematical manipulations obtained 53.13%, examining the validity of arguments obtained 54.69%, and drawing logical conclusions obtained 65.63%. This pattern indicates that students' initial abilities were relatively lower in activities requiring conjecture formation and justification than in activities involving drawing conclusions. This condition was an important consideration in designing the game because each indicator was subsequently assigned to a different game level and integrated with the problem-solving stages.

The integration of the five indicators with Polya's four stages provided a structure for the thinking process within the game. The activity of making conjectures was placed within the processes of understanding information and planning the solution. Mathematical manipulation was facilitated when students carried out their selected strategies. Constructing proofs or providing reasons emerged when students were required to explain the basis of their solutions. Examination of the validity of arguments was facilitated through the looking-back stage, while drawing conclusions was directed toward formulating the final result based on the solution process. Thus, Polya's stages in this study functioned as the structure of the problem-solving process, while the five mathematical reasoning indicators represented the abilities developed through that process. Scaffolding through hints and feedback provided additional support when students experienced difficulties in following the process. The assistance was provided as guidance during problem solving so that students remained directed toward finding the solution rather than simply receiving the final answer. This mechanism is consistent with Saputra et al. (2024) and Hartiningrum et al. (2024), who explained that gradual assistance can support students in developing understanding and mathematical thinking abilities. Therefore, the hint and feedback features in this study served a pedagogical function as forms of scaffolding, rather than merely functioning as game features.

These findings suggest that the main contribution of the developed game does not lie solely in the use of Scratch but also in the organization of reasoning activities within the game structure. Scratch provided an interactive learning environment, Polya's stages organized the problem-solving process, and the five reasoning indicators determined the abilities developed at each level. This integration provides an alternative learning-media design that connects gameplay activities with mathematical thinking processes. The findings reinforce previous research regarding the potential of Scratch and problem solving in mathematics learning, while also showing that the improvement obtained in this study remained in the moderate category. Therefore, the results are more appropriately interpreted as an indication that the use of the game has the potential to support the development of mathematical reasoning ability rather than as evidence of a causal effect. Future research should employ experimental or quasi-experimental designs with comparison groups to provide stronger evidence regarding the contribution of the game to changes in students' mathematical reasoning ability.

Kognitif: Jurnal Riset HOTS Pendidikan Matematika
https://doi.org/10.51574/kognitif.v6i3.5637 Volume 6, No 3, July - September 2026, pp. 1256 –1273

Conclusion

This study produced a problem-solving-based mathematics educational game using Scratch on statistics material for tenth-grade high school students. The validation results showed that the game was categorized as very valid, while the practicality test results indicated that the game was categorized as practical. The main field trial showed an increase in the mean mathematical reasoning score from 51.56 on the *pretest* to 79.19 on the *posttest*, with an N-Gain value of 0.567, categorized as moderate. These results indicate that students' mathematical reasoning scores increased after using the game within the context of this study. Since the study did not include a comparison group, the improvement cannot be used to establish a causal effect of the game on mathematical reasoning ability. This study has limitations related to the implementation of the field trial at only one school and with a limited number of students. These conditions limit the generalizability of the findings to broader student characteristics and school contexts. Future studies are recommended to employ experimental or quasi-experimental designs with comparison groups and involve participants from more than one school to obtain stronger evidence regarding the effect of using the game on students' mathematical reasoning ability.

Conflict of Interest

The authors declare that there is no conflict of interest related to the publication of this article.

Auhor Contributions

N.A. contributed to conceptualization, methodology, investigation, data collection, data analysis, development of the educational game, and writing the original manuscript. N. contributed to supervision, methodological guidance, review, and editing of the manuscript. M.F. contributed to supervision, academic guidance, review, and editing of the manuscript. All authors read and approved the final version of the manuscript. The percentage contributions of each author were N.A.: 70%, N.: 15%, and M.F.: 15%.

Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

References

Agustin, S. S., Purwanto, S. E., Ma'arif, S., & Soebagyo, J. (2021). Analisis Kemampuan Penalaran Matematis Peserta Didik dengan Penyajian Masalah Open-Ended Pada Pembelajaran Daring. *Jurnal Riset HOTS Pendidikan Matematika*, *1*(1), 66–80. https://doi.org/10.51574/kognitif.v1i1.19

Anggraini, R. D., Murni, A., & Gunawan, T. (2022). Pengembangan Perangkat Pembelajaran Menggunakan Problem Based Learning Pada Materi SPLTV Untuk Memfasilitasi Kemampuan Pemecahan Masalah Matematis. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, *06*(03), 3136–3147. https://doi.org/10.31004/cendekia.v6i3.1658

Ardiyanti, D., & Tyas, D. N. (2025). Pengembangan Multimedia Pembelajaran Interaktif Shape Safari Berbasis Scratch untuk Meningkatkan Hasil Belajar Matematika Materi

Karakteristik Bangun Datar. *Jurnal Pendidikan MIPA*, *15*(1), 2621–9166. https://doi.org/10.37630/jpm.v15i1.2575

Branch, R. M. (2009). *Instructional Design: The ADDIE Approach*. https://doi.org/10.1007/978-0-387-09506-6

Efendi, T. N., Kartini, & Anggraini, R. D. (2024). Pengembangan Instrumen Tes Kemampuan Penalaran Matematis pada Materi Barisan dan Deret Kelas XI SMA/MA. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, *08*(01), 811–826. https://doi.org/10.31004/cendekia.v8i1.2650

Faradilla, Y., Pamungkas, M. D., & Rahmawati, F. (2023). Pengembangan Multimedia Pembelajaran "Scraperat" Untuk Meningkatkan Kemampuan Penalaran Matematis Siswa Kelas IX. *RANGE: Jurnal Pendidikan Matematika*, *5*(1), 50–62. https://doi.org/10.32938/jpm.v5i1.4631

Fitriyah, I. M., Putro, N. H. P. S., & Apino, E. (2022). Meta Analysis Study: Effectiveness of Problem Based Learning on Indonesian Students' Mathematical Reasoning Ability. *Jurnal Riset Pendidikan Matematika*, *9*(1), 36–45. https://doi.org/10.21831/jrpm.v9i1.46447

Fortuna, I. D., Yuhana, Y., & Novaliyosi. (2021). Pengembangan Lembar Kerja Peserta Didik dengan Problem Based Learning untuk Kemampuan Berpikir Tingkat Tinggi. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, *05*(02), 1308–1321. https://doi.org/10.31004/cendekia.v5i2.617

Hartiningrum, E. S. N., & Wahab, A. (2024). Kajian Teori: Pandangan Vygotsky dalam Pembelajaran Matematika Menggunakan Problem Posing Di Kurikulum Merdeka. *Jurnal Ilmiah Soulmath : Jurnal Edukasi Pendidikan Matematika*, *12*(2), 107–120. http://doi.org/10.25139/smj.v12i2.7822

Indah Widiyanti, S., & Desniarti. (2025). Pengembangan LKPD dengan Pendekatan Problem Based Learning Berbantuan Scratch untuk Meningkatkan Keterampilan Pemecahan Masalah Matematis Siswa. *Jurnal Ilmiah Pendidikan Dasar*, *10*(03), 481–488. https://doi.org/10.23969/jp.v10i03.30092

Indahwati, R., Basri, H., & Aini, S. D. (2024). Pengembangan Media Pembelajaran Berbasis Scratch Elemen Bilangan Fase A pada Kurikulum Merdeka. *Jurnal Riset HOTS Pendidikan Matematika*, *4*(4), 1626–1637. https://doi.org/10.51574/kognitif.v4i4.2632

Mariza, E., Karimah, N. I., Laelasari, L., & Putri, D. P. (2022). Kesulitan Belajar Peserta Didik dalam Menyelesaikan Soal Cerita SPLTV Berbasis Pemecahan Masalah. *Jurnal Karya Pendidikan Matematika*, *9*(2), 45. https://doi.org/10.26714/jkpm.9.2.2022.45-52

Maulidina Rahadi, N., & Ismiyati, N. (2024). Pengembangan Media Pembelajaran Matematika Berbantuan Aplikasi Scratch Siswa Kelas VII SMPN 7 Balikpapan Tahun Akademik 2023/2024. *De Fermat : Jurnal Pendidikan Matematika*, *7*(2), 83–91. https://doi.org/10.36277/defermat.v7i2.335

Monalisa, L. A., Kristiana, A. I., Fatahillah, A., Wihardjo, E., Aulia, V. A., Hussen, S., & Falebita, O. S. (2025). Development of Interactive Learning Media Assisted by Scratch to Enhance Understanding of Rectangle and Square Concepts. *Research and Development in Education (RaDEn)*, *5*(1), 741–753. https://doi.org/10.22219/raden.v5i1.4 0008

National Council of Teachers of Mathematics. (2000). Principles and Standards for School Mathematics. In National Council of Teachers of Mathematics. Reston, VA: National Council of Teachers of Mathematics.

Novaliyosi, Yuhana, Y., & Khaerunnisa, E. (2022). Pengembangan Lembar Kerja Interaktif untuk Mendukung Kemampuan Berpikir Tingkat Tinggi Pada Pembelajaran Daring. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *11*(3), 1640–1655. https://doi.org/10.24127/ajpm.v11i3.5334

Kognitif: Jurnal Riset HOTS Pendidikan Matematika
https://doi.org/10.51574/kognitif.v6i3.5637 Volume 6, No 3, July - September 2026, pp. 1256 –1273

Nursalam, Angriani, A. D., & Usman, H. (2017). Pengembangan Tes Kemampuan Penalaran Matematis Peserta Didik Madrasah Tsanawiyah di Makasar. *Lentera Pendidikan : Jurnal Ilmu Tarbiyah Dan Keguruan*, *20*(1), 85–97. https://doi.org/10.24252/lp.2017v20n1a7

Prastika, V. Y. A., Riyadi, & Siswanto. (2021). Mathematical Reasoning Ability of Junior High School Viewed From Logical Mathematical Intelligence. *Journal of Physics: Conference Series*, *1918*(4). https://doi.org/10.1088/1742-6596/1918/4/042067

Putra, D. D., Fathurrohman, M., Nindiasari, H., & Novaliyosi, N. (2025). Penggunaan Labirin; the Board Game Terhadap Kemampuan Pemecahan Masalah Matematis Siswa Smk Ditinjau Dari Kemampuan Awal Matematika. *Pedagogy: Jurnal Pendidikan Matematika*, *10*(4), 1835–1853. https://doi.org/10.30605/pedagogy.v10i4.7240

Putra, D. D., Novaliyosi, N., Nindiasari, H., & Fathurrohman, M. (2025). Pengembangan Instrumen Tes Kemampuan Pemecahan Masalah Matematis Siswa Sekolah Menengah Kejuruan. *Edukasiana: Jurnal Inovasi Pendidikan*, *4*(3), 967–983. https://doi.org/10.56916/ejip.v4i3.1610

Rahmania, M. D., Fatah, A., & Anriani, N. (2023). Pengembangan Media Pembelajaran Interaktif Matematika Berbasis Web Articulate Storyline untuk Meningkatkan Minat Belajar Siswa SMP. *Jurnal Absis: Jurnal Pendidikan Matematika Dan Matematika*, *5*(2), 653–665. https://doi.org/10.30606/absis.v5i2.1777

Salamah, F., & Sudihartinih, E. (2024). Pengembangan Multimedia Pembelajaran Matematika Pada Materi Grafik Persamaan Garis Lurus Berbasis Android Menggunakan Aplikasi Scratch. *Journal Mathematics Education Sigma [JMES]*, 141–152. https://doi.org/10.30596/jmes.v5i2.19639

Saputra, R., Novaliyosi, Syamsuri, & Hendrayana, A. (2024). Systematic Literature Review: Strategi Scaffolding dalam Pembelajaran Matematika Untuk Meningkatkan Pemahaman Siswa. *Jurnal Cendekia : Jurnal Pendidikan Matematika*, *8*(2), 1697–1710. https://doi.org/10.31004/cendekia.v8i2.3312

Son, A. L. (2019). Instrumentasi Kemampuan Pemecahan Masalah Matematis: Analisis Reliabilitas, Validitas, Tingkat Kesukaran dan Daya Beda Butir Soal. *Gema Wiralodra*, *10*(1), 41–52. https://doi.org/10.31943/gemawiralodra.v10i1.8

Suryanti, Pramesti, C., & Sidik, R. S. R. (2022). Kesalahan Penalaran Matematis Pada Materi Persamaan Diferensial. *Journal Numery*, *9*(1), 14–26. https://doi.org/10.46244/numeracy.v9i1.1755

Syarifah, N. R., Yuhana, Y., & Fathurrohman, M. (2025). Development of Scratch-Based Mathematics Learning Media To Increase Students' Learning Motivation. *EMTEKA: Jurnal Pendidikan Matematika*, *6*(2), 920–934. https://doi.org/10.24127/emteka.v6i1.8399

Authors Biography

Nisa Aghnia is an undergraduate student in the Department of Mathematics Education, Faculty of Teacher Training and Education, Universitas Sultan Ageng Tirtayasa, Banten, Indonesia. Her research interests include mathematics education, educational game development, *problem solving*, mathematical reasoning, and technology-based learning.
✉ nisaaghnia705@gmail.com

	**Novaliyosi** is a lecturer and researcher in the Department of Mathematics Education, Faculty of Teacher Training and Education, Universitas Sultan Ageng Tirtayasa, Banten, Indonesia. Her research interests include mathematics education, *higher-order thinking skills* (HOTS), mathematical reasoning, mathematical *problem solving*, and mathematics learning media development. ✉ novaliyosi@untirta.ac.id
	**Maman Fathurrohman** is a lecturer and researcher in the Department of Mathematics Education, Faculty of Teacher Training and Education, Universitas Sultan Ageng Tirtayasa, Banten, Indonesia. His research interests include mathematics education, educational game development, game-based learning, mathematics learning media, educational technology, and ICT integration in mathematics education. ✉ mamanf@untirta.ac.id