

Analysis of Islamic boarding school students' higher-order thinking skills in number topics based on Bloom's revised taxonomy

Mufarrido Husnah*, Jarnawi Afgani Dahlan, Dadan Dasari, Tatang Herman

Universitas Pendidikan Indonesia, West Java, Indonesia

*Correspondence: mufarrido.husnah@upi.edu

Received: 05 January 2025 / Accepted: 05 February 2026 / Published Online: 12 June 2026
© The Author(s) 2026

Abstract

This study aimed to analyze the Higher Order Thinking Skills (HOTS) of Islamic boarding school students in number topics, specifically number factorization, Least Common Multiple (LCM), and Greatest Common Factor (GCF), based on Anderson's cognitive taxonomy. The study employed a qualitative descriptive approach, using three questions designed to measure higher-order thinking skills at the analysis (C4), evaluation (C5), and creation (C6) levels. The research sample consisted of 22 Islamic boarding school students who were selected using stratified random sampling. The results of the analysis showed that students' abilities in the aspects of analysis and evaluation were quite good, while their abilities in the aspect of creation were still low. Most students were categorized as having moderate and low levels of ability based on the scores obtained. These findings indicated the need for more focused and innovative learning strategies to promote the development of students' HOTS, especially in the context of Islamic boarding schools, which had limited time allocated for general subjects. This study recommended the integration of more challenging and creative learning approaches to improve the quality of students' higher-order thinking skills.

Keywords: Bloom's Revised Taxonomy, Higher-Order Thinking Skills, Islamic Boarding Schools, Number Factorization, Qualitative Study

How to Cite: Husnah, M., Dahlan, J.A., Dasari, D., & Herman, T. (2026). Analysis of higher-order thinking skills of Islamic boarding school students on number material based on Bloom's revised taxonomy. *AXIOM : Jurnal Pendidikan dan Matematika*, 15(1), 15-27. <https://doi.org/10.30821/axiom.v15i1.23059>

Introduction

Mathematics is the science or knowledge of learning or logical thinking that is essential for humans to live and underlies the development of modern technology. Mathematics itself is not just a collection of abstract concepts but also a logical system that serves to organize ideas and understand the world around us. In addition to its theoretical aspects, mathematics has real-world applications in everyday life, making it a potential tool for developing students' higher-order thinking skills (Thanheiser, 2023). Based on Indonesia's Merdeka Curriculum Learning Outcomes, mathematical ability is a basic skill that children need to understand the world and use that ability in their daily activities.

Memorizing and understanding the material alone is not enough to achieve the objectives of mathematics learning (Brookhart, 2010). Students' potential needs to be improved in their ability to analyze, evaluate, and create in order to achieve the objectives of mathematics



learning comprehensively (Ismail et al., 2024). In the cognitive taxonomy, these three abilities are included in *Higher Order Thinking Skills* (HOTS). In the revised taxonomy by (Anderson & Krathwohl, 2001), HOTS covers the three highest cognitive levels, namely C4 (analyzing), C5 (evaluating), and C6 (creating). At level C4 (analyzing), students are expected to be able to break down information into smaller parts, understand the relationships between these parts, and recognize existing structures or patterns. For example, students can analyze mathematical data to identify certain trends or patterns. Furthermore, at level C5 (evaluating), students demonstrate the ability to assess information, arguments, or solutions based on predetermined criteria. Evaluation as a higher-order thinking skill includes the ability to distinguish essential information or core concepts from things that are only superficially interesting, as well as making critical judgments about ideas and concepts that are read or studied (Piltan, 2010). Finally, at the highest level, C6 (creating), students are required to synthesize various elements of information, both new and existing, to form new ideas, new ways of thinking, or create innovative solutions. Examples include when students create mathematical models or design new approaches to solving complex problems.

Higher Order Thinking Skills (HOTS) have become a significant topic of discussion and continue to be developed among educators in Indonesia in recent years. The government realizes that the education system still needs to be strengthened in the application of HOTS as one of the internationally recognized standards in Indonesia (Kunanti, 2020). Higher-order thinking skills are one of the key abilities that need to be developed in mathematics learning, even in group learning (Vijayaratnam, 2012). HOTS includes the ability to analyze, evaluate, and create, which can be measured through formative and summative assessments (Manik et al., 2020). These skills not only help students to understand concepts in depth but also train them to think critically, creatively, and produce innovative solution (Nenohai et al., 2024). By having good HOTS abilities and a positive attitude, students can solve problems effectively, make the right decisions, and demonstrate good character (Rahmi & Azrul, 2022).

The urgency of mastering HOTS is increasing in the 21st century due to the challenges of modern life that require individuals to be able to deal with complex problems, think critically, and adapt to change (Yaniawati et al., 2021). These skills are an important indicator for preparing students to face globalization, technological advances, and an increasingly dynamic world of work. HOTS enables students to not only absorb information or facts (Nowlan et al., 2023) but also to process, evaluate, and create something new based on their knowledge (Nenohai et al., 2024). According to (Brookhart, 2010), higher-order thinking skills in classroom learning can be classified into three main categories. First, these skills are defined as the ability to transfer knowledge to various contexts. Second, higher-order thinking skills are understood as critical thinking skills. Third, these skills include the ability to solve problems.

The development of HOTS in mathematics learning in Islamic boarding schools faces a number of challenges. The low level of higher order thinking skills (HOTS) among Indonesian students, especially those in Islamic boarding schools, has prompted various studies related to this aspect. Research by (Purbaningrum, 2017) shows that junior high school students' HOTS abilities are still relatively low when viewed from visual, auditory, and kinesthetic learning styles. Similar results were also found by (Milenia et al., 2022), who stated that junior high school students' abilities in solving HOTS-based math problems were lowest at the creating

(C6) level. Students with high abilities can solve HOTS problems completely, while students with low abilities only complete some stages of problem solving.

In the context of HOTS assessment, research by Fanani (2018) shows that assessments that measure the ability to analyze (C4), evaluate (C5), and create (C6) can increase student learning motivation (Brookhart, 2010). However, major challenges in mathematics learning in Islamic boarding schools include ontogenetic, didactic, and epistemological barriers (Ramdhani et al., 2021). A dense curriculum and time constraints hinder the learning process and have an impact on decreasing student motivation. (Fathani, 2019) also emphasizes the importance of mathematics learning in Islamic boarding schools to equip students with logical, analytical, and creative thinking skills, even though the approach currently used is not yet optimal.

Another factor that affects low HOTS abilities is a lack of motivation and practice in solving HOTS questions (Fahri & Setyaningrum, 2023). To overcome this problem, a more effective learning approach is needed. (Karwadi et al., 2024) in their meta-analysis showed that the Active Learning method is effective in improving HOTS, with the blended learning method showing the most significant impact. In addition, the development of Android-based mobile learning applications has also been proven to improve junior high school students' HOTS and motivation to learn mathematics (Yaniawati et al., 2021). Contextual approaches also yield positive results. For example, Javanese vegetable salad context-based mathematics learning is effective in improving students' HOTS (Gembong et al., 2022). At the university level, synchronous and asynchronous blended learning discussion methods help improve students' HOTS through information sharing, idea exploration, and critical assessment (Rahmi & Azrul, 2022).

However, these studies were generally conducted in public schools or universities that have full-time allocations for general subjects and broad access to learning technology. To date, there has been very limited research specifically exploring students' Higher Order Thinking Skills (HOTS) in Islamic boarding schools, especially those with limited time for general subjects such as mathematics. Nevertheless, there are several studies conducted in Islamic boarding schools with integrated education systems, such as that conducted by (Hanafi et al., 2021), although the focus of the study is not directly related to the development of HOTS. Therefore, this study attempts to fill this gap by examining students' higher-order thinking skills in integrated Islamic boarding schools more specifically based on Anderson's cognitive taxonomy.

By understanding the challenges and needs of HOTS development in Islamic boarding schools, this study is expected to make a real contribution to improving the quality of mathematics learning. The focus of this study is to analyze the higher-order thinking skills of Islamic boarding school students in number-related material based on Bloom's Revised Taxonomy. This study also aims to provide a comprehensive overview of students' abilities in analyzing, evaluating, and creating. The results are expected to not only serve as a basis for developing more effective learning strategies, but also support the integration of 21st-century skills into the Islamic boarding school education system.

Method

This research uses a qualitative research method. According to (Creswell & Creswell, 2018) qualitative research is an approach used to explore and understand the meanings that emerge from the experiences of individuals or groups related to a social issue. The description is of students' higher-order thinking skills in numerical material based on Bloom's Revised Taxonomy cognitive domain. The description was obtained through a written test conducted in December 2024. According to (Sugiyono, 2013), qualitative research is an approach based on postpositivism philosophy and is used to examine phenomena in natural conditions. In this approach, the researcher acts as the main instrument, using data collection techniques that combine various sources (triangulation) and applying inductive data analysis.

This study involved 22 students from a private Islamic junior high school in Sukabumi. Participants were selected purposively, taking into account the diversity of students' academic abilities, which included high, moderate, and low abilities. This diversity was not limited to one specific ability, but was intended to represent the variation in students' abilities in general. Thus, the data obtained was expected to provide a more comprehensive picture in accordance with the research objectives. The instrument used in this study was HOTS questions on numbers, specifically number factorization, which had been validated by an expert lecturer. The validation process included examining the suitability of higher-order thinking ability indicators, linguistic accuracy, and the clarity of the question context. After the instruments were validated, the test sheet consisted of three questions designed to measure students' cognitive abilities based on Bloom's Revised Taxonomy at the analysis (C4), evaluation (C5), and creation (C6) levels. The first question (C4) measures analytical skills by asking students to determine the number of groups containing the same number of objects without remainder and to calculate the number of both types of objects in each group accurately. The second question (C5) measures evaluation skills by asking students to check the accuracy of the GCF of the two given numbers and to provide reasons for their answers. The third question (C6) measures creativity, where students are asked to compose a story question related to daily activities involving the LCM.

This instrument was developed based on Bloom's Revised Taxonomy. This taxonomy is a revision of Bloom's taxonomy published in 2001 by Anderson and Krathwohl. The main difference between the revised taxonomy and the original version lies in the addition of two dimensions, namely the knowledge dimension and the cognitive process dimension (Brookhart, 2010). The knowledge dimension classifies the types of knowledge that students must master into four categories, namely factual, conceptual, procedural, and metacognitive knowledge. Factual knowledge includes basic information such as terms and symbols used in lessons. Conceptual knowledge refers to an understanding of the relationships between concepts, principles, or theory. Procedural knowledge involves understanding the methods or steps for completing a task or solving a problem, while metacognitive knowledge relates to students' awareness of their thinking processes and learning strategies. Meanwhile, the cognitive process dimension consists of six levels of thinking, namely remembering (C1), understanding (C2), applying (C3), analyzing (C4), evaluating (C5), and creating (C6). The main difference with Bloom's Taxonomy lies in the last two levels, where the order of "evaluating" and "creating" is reversed. In the Revised Bloom's Taxonomy, creating is placed at the highest level because it

is considered to represent the most complex thinking ability, namely when students are able to produce something new based on deep understanding and analysis. This framework allows educators to be more careful in designing questions, assessing learning outcomes, and understanding students' thinking abilities at various cognitive levels.

There are three questions that will be tested in essay form. The scores given are determined based on predetermined assessment guidelines with a score range of 0 to 15 per question, so the total score for all questions is 45. The scores from the students' answers will then be converted into test scores using the following Formula (1):

$$Test\ score = \frac{scores\ obtained}{maximum\ score} \times 100 \quad (1)$$

The test scores of students will be calculated to determine their average score. The scores obtained will then be grouped into three groups based on Table 1 (Fahri & Setyaningrum, 2023).

Table 1. High-Level Thinking Ability Groups

Score	Group
76-100	High
51-75	Moderate
0-50	Low

Furthermore, to determine the ability of the three groups of students in answering HOTS questions, a cognitive domain analysis was conducted at the analyzing (C4), evaluating (C5), and creating (C6) levels. The students' answers at each cognitive level were analyzed based on the relevant knowledge dimensions according to the indicators, namely procedural knowledge in the analyzing level (C4) questions, conceptual knowledge in the evaluating level (C5) questions, and metacognitive knowledge in the creating level (C6) questions. The indicators for the number factorization material questions, which include the knowledge dimension and cognitive level dimension, are provided in Table 2.

Table 2. HOTS Question Indicators for Number Factorization Material based on Bloom's Revised Taxonomy

Question Level	Dimension Knowledge	Question Indicators
Analyzing	Procedural	Students are able to analyze the distribution of books and pencils using the GCF to determine the correct solution. Students explain the steps of the solution systematically until they find the maximum number of students and the number of books and pencils per student.
Evaluating	Conceptual	Students are able to evaluate the truth of a statement related to the GCD of two numbers. Students use the concept of GCF correctly, including comparing the calculation results with the given statement, and providing reasons to support their decisions.
Creating	Metacognitive	Students are able to design story problems related to the LCM based on observations or daily experiences. Students demonstrate awareness of context and logic in constructing problems.

The analysis of student responses in this study was conducted using content analysis. The analysis process began with a thorough reading of all student responses, followed by the identification of the suitability of the responses to the competency indicators at cognitive levels

C4, C5, and C6. Next, student responses were classified based on reasoning characteristics, solution strategies, and errors that appeared in each question item. This classification became the basis for interpreting students' reasoning patterns at each cognitive level. To ensure the validity and consistency of the analysis, peer checking was carried out by the supervising lecturer to review the classification and interpretation process that had been carried out.

Results

To measure students' higher-order thinking skills, the researcher used an instrument in the form of HOTS essay questions. The instrument consisted of three questions on numbers, covering number factorization, GCD, and LCM. Based on the results of the test, the data obtained in the form of test scores is presented in Table 3.

Table 3. Test Results for HOTS-Based Questions Based on Question Level

Question Level	Average Score
Analyzing (C4)	62,72
Evaluating (C5)	40
Creating (C6)	54,84

The results show variations in average scores across three question levels. The analysis question level (C4) had the highest average score of 62.72, indicating that students tended to be more capable of solving problems that require analytical skills. Conversely, the evaluation level (C5) had the lowest average score of 40, indicating that students' ability to evaluate information or arguments still needs to be improved. Meanwhile, at the creating level (C6), the average score of 54.84 shows that students' ability to develop new ideas or create solutions is at a moderate level, although still below the average for analytical questions. Meanwhile, the categorization of subjects into high, moderate, and low abilities in solving HOTS-based questions can be seen in Table 4. Based on the results in Table 4, most subjects were in the low category, namely 11 people (50%). This indicates that half of the subjects have HOTS abilities that still need to be improved. Meanwhile, the moderate category includes 7 people (31.8%), indicating that one-third of the subjects have adequate HOTS abilities. On the other hand, only 4 people (18.2%) are in the high category, reflecting that HOTS abilities in subjects with the best performance are still relatively few. Next, the results of the analysis based on Bloom's Revised Taxonomy for C4, C5, and C6 questions are presented, along with an interpretation of the variations in answers across each subject category.

Table 4. Results of the HOTS-Based Problem-Solving Test by Subject Category

Category	Frequency	Percentage
High	4	18,2%
Moderate	7	31,8%
Low	11	50%

Data analysis was conducted on three categories of subjects based on test scores, namely high, moderate, and low categories. For each ability category, the researcher selected one sample answer from each cognitive level (C4, C5, and C6). The selection was made purposively by considering the answers that were most representative of the characteristics of the category and the highest suitability with the ability indicators at the level being analyzed. The selection of representative samples aimed to present a comprehensive picture of the

characteristics of student reasoning in each ability category. The discussion of the results at each HOTS cognitive level, accompanied by a sample of student answers that had been selected purposively, is presented in the following section.

Analyzing Level (C4)

At the analyzing level (C4), students were asked to analyze problems in contextual questions to determine the appropriate solution steps, namely using the GCF concept. The knowledge dimension observed was procedural knowledge, which involved students' ability to apply number factorization procedures and select the appropriate procedure to determine the largest number of students. To provide a clearer picture of student achievement at the analyzing level (C4), the test results were then grouped based on high, moderate, and low ability categories. The test results for each subject category on question C4 are shown in Table 5.

Table 5. Test Results Based on Subject Categories on Question C4

Category	Student Responses	Translation of Student's Response
High		<p>GCD = $2 \times 5 = 10$ (representing the number of students)</p> <p>40 : 10 = 4 books per student 50 : 10 = 5 books per student</p>
Moderate		<p>Given: 40 books and 50 pencils Asked: How many books and pencils are received by each student Answer: LCM = $2 \times 5 \times 4 = 200$ GCD = $2 \times 5 = 10$ There are 10 students, and each student receives 4 books and 5 pencils.</p>
Low		<p>KPK : Least Common Multiple (LCM) FPB : Greatest Common Divisor (GCD)</p>

In the initial stage, all categories of students were able to break down the information correctly, i.e., separating the number of books and pencils given. Next, the three categories of students used number factorization procedures in various ways but produced correct answers. At the stage of determining the largest number of students, students in the high category immediately chose to use the GCF method, demonstrating a deeper understanding of the concept of analysis. Meanwhile, students in the moderate and low categories first compared the results of the LCM and GCF. Students in the moderate category finally used the GCF to solve the problem, while students in the low category did not continue the process to the end,

so they could not conclude the results. After finding the largest number of students using the GCD, the high and moderate category students were able to determine the number of books and pencils received by each student. However, the moderate category students did not clearly list the division steps, so the procedure used could not be further analyzed.

From this analysis, it can be concluded that high-level students demonstrated more structured and direct analytical skills in choosing the appropriate solution method. Intermediate-level students were able to achieve the correct final result despite their less detailed steps, while low-level students showed difficulty in completing the final stage of solving the problem. This indicates differences in the level of understanding and analytical skills between student categories at the C4 level.

Evaluating Level (C5)

At the evaluating level (C5), students are asked to determine the truth of a statement and provide reasons for their decisions. The dimension of knowledge observed is conceptual knowledge, because students must understand the concept of GCF and evaluate the truth of the statement based on evidence. The evidence provided by students is in the form of a factorization table of the two numbers given. The statement in the question given is a false statement. Students are asked to check the truth of the statement and provide their reasons.

Table 6. Test Results Based on Subject Categories in Question C5

Category	Student Responses	Translation of Student's Response
High		<p>$GCD = 3 \times 5^2 = 75$ Because to find the greatest common divisor of 225 and 375, the factors are $3 \times 5 \times 5$, so $3 \times 5 = 15 \times 5 = 75$. Therefore, the result of the problem is 75.</p>
Moderate		<p>$GCD = 5 \times 5 = 25$ Reason: how I found it is that the numbers are divided by the same number repeatedly (that is, the numbers are divided by the same factor continuously, so I looked for a number that can divide all the numbers involved)</p>
Low		<p>KPK : Least Common Multiple (LCM)</p>

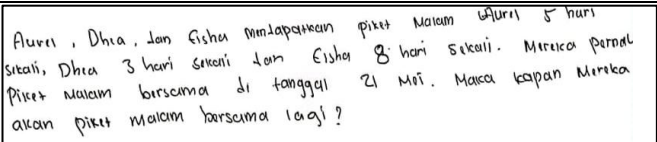
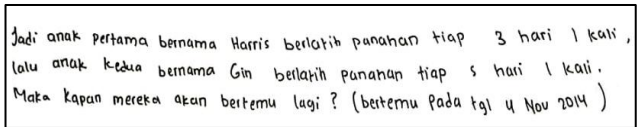
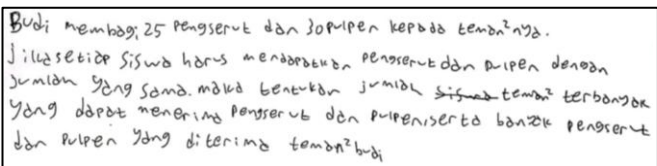
Based on the results in Table 5, the three categories of students showed variations in evaluating the truth of the given statements. Students in the high category used the GCF concept correctly, including comparing the calculation results with the given statements and providing reasons to support their conclusions. Students in the moderate category had difficulties in the factorization process, which resulted in incorrect GCF results. Even though the GCF results were incorrect, students in the moderate category still provided reasons that showed their efforts to evaluate the statements. Meanwhile, students in the low category did not succeed in evaluating the GCF of the two numbers correctly and did not present conclusions or supporting reasons.

From this analysis, it can be concluded that the success of students in evaluating the truth of statements at the C5 level is greatly influenced by their mastery of conceptual knowledge. High-level students demonstrate a good understanding of the GCF concept and are able to evaluate the truth of statements based on valid evidence. Conversely, students in the middle and low categories faced difficulties in applying conceptual knowledge, which affected the accuracy of their evaluation results and the reasons they gave. These results show the importance of conceptual knowledge mastery in solving problems at the analyzing level (C5).

Creating Level (C6)

At the creating level (C6), students were asked to create story problems that were relevant to everyday life and related to GCD. This activity involved metacognitive knowledge, as students had to think creatively (Heong et al., 2012) while reflecting on how to connect the concept of GCD to real-life situations. Students were asked to compose a logical and clear problem.

Table 7. Test Results Based on Subject Categories in C6 Questions

Category	Student Responses	Translation of Student's Response
High		<p>Aurel, Dhea, and Eisha have night duty on different schedules. Aurel has night duty every 5 days, Dhea every 3 days, and Eisha every 8 days. They had night duty together on the 21st May. So, when will they have night duty together again?</p>
Moderate		<p>The first child named Harris practices archery every 3 days. The second child named Gin practices archery every 5 days. So, when will they meet again? (They meet again on November 4, 2014.)</p>
Low		<p>Budi distributes 25 erasers and 30 pens to his friends. If each student must receive the same number of erasers and pens, determine the maximum number of friends who can receive them,</p>

as well as the number of erasers
and pens received by each friend.

The analysis results show that students in the high and moderate categories are able to compose questions that are in line with the GCD concept and demonstrate an awareness of context and logic in their composition. Students in the high category use concise but clear language so that the questions are easy to understand, while students in the moderate category compose questions using concise language, but their presentation is less structured because some information is separated from the question sentence. Students in the low category were also able to compose questions using clear language and complete information, but they used the wrong concept, namely GCF instead of LCM, so their questions did not match the instructions. Overall, all students showed creativity by relating the questions to the context of daily life in the Islamic boarding school environment, although there were differences in the level of clarity of language and conceptual appropriateness among the three categories of students.

Discussion

The results of the study indicate that the higher-order thinking skills (HOTS) of Islamic boarding school students in numeracy vary at each level of Bloom's Revised Taxonomy, namely analyzing (C4), evaluating (C5), and creating (C6). In general, students in the high category show more consistent abilities in solving HOTS questions, especially at the C6 level, compared to students in the moderate and low categories. These findings support the research by (Milenia et al., 2022), which found that students with high abilities are able to complete HOTS questions completely, while students with low abilities tend to have difficulties, especially at the creating level.

Theoretically, this study contributes to strengthening the understanding of differences in HOTS development among student ability categories, particularly in the context of Islamic boarding schools. These findings also support (Brookhart, 2010) concept of the ability to transfer mathematical knowledge to everyday contexts, while confirming the epistemological barriers identified by (Ramdhani et al., 2021), where low-category students still have difficulty applying concepts appropriately. Thus, this study adds empirical evidence regarding the relationship between students' mathematical abilities and the learning context in Islamic boarding schools.

From a practical perspective, these findings emphasize the need for contextual learning designs oriented toward the development of HOTS, especially to improve students' analytical, evaluative, and creative abilities. For example, although students showed creativity in composing KPK-based story questions, limited learning time and a focus on the religious curriculum were factors that limited the optimal development of HOTS. Therefore, teachers and curriculum developers need to design learning strategies that take into account differences in student abilities so that all categories can develop to their full potential.

In addition, this study supports the findings (Fathani, 2019) which emphasize the need for a learning approach that equips Islamic boarding school students with analytical and creative thinking skills. Although Islamic boarding school students showed creativity in composing KPK-based story questions, limited learning time and focus on the religious curriculum were

factors that hindered the optimal development of HOTS. The findings of this study make an important contribution to the development of mathematics learning in Islamic boarding schools, particularly in integrating HOTS into a learning context that is relevant to the daily lives of students. The practical implications include the need for contextual learning designs to improve students' analytical, evaluative, and creative abilities. In addition, this study provides an initial overview of the obstacles and opportunities in developing HOTS in the pesantren environment, which can serve as a basis for further research.

This study has several limitations, such as focusing only on numerical material and data collection limited to one Islamic boarding school. The implication of these limitations is that the results of the study may not fully represent the population of Islamic boarding school students in general. Therefore, further research can expand the scope of material and the number of Islamic boarding schools to provide more accurate and representative results.

Conclusion

This study shows that Islamic boarding school students' higher-order thinking skills (HOTS) in numerical material vary at each level of Bloom's Revised Taxonomy. Students in the high category tend to be able to solve HOTS problems accurately and consistently, especially at the creating level (C6), compared to students in the moderate and low categories. These findings emphasize the importance of a contextual and HOTS-oriented learning approaches to improve students' analytical, evaluative, and creative skills in Islamic boarding schools.

The practical implications of this study include the need to develop learning strategies that are relevant to the daily lives of Islamic boarding school students, such as contextual learning design and the integration of technology-based methods to improve HOTS mastery. Further research is recommended to expand the scope of the material, increase the number of Islamic boarding school samples, and integrate learning innovations to provide a more representative picture and support the development of HOTS as a whole.

Declarations

- Author Contribution : MH : Formal Analysis, Writing - Original Draft, Editing, Methodology.
JAD : Conceptualization, Validation and Supervision.
DD : Supervision.
TH : Supervision.
- Funding Statement : No funding.
- Conflict of Interest : The authors declare no conflict of interest.
- Additional Information : No additional information is available for this paper.

References

- Anderson, L.W., & Krathwohl, D.R. (2001). *A taxonomy for learning, teaching, and assessing: a revision of bloom's*. New York: Longman.
- Brookhart, S.M. (2010). *How to assess higher-order thinking skills in your classroom*. United State: ASCD.

- Creswell, J.W., & Creswell, J.D. (2018). Research design: qualitative, quantitative and mixed methods approaches. In *SAGE Open* (5th ed.). Yogyakarta: Pustaka Belajar.
- Fahri, J., & Setyaningrum, V. (2023). Analisis kemampuan berpikir tingkat tinggi peserta didik materi volume bangun ruang. *J-P*, 5(1), 743–754. <https://doi.org/10.31932/j-pimat.v5i1.2392>
- Fanani, M.Z. Strategi pengembangan soal HOTS pada Kurikulum 2013. *Edudeena : Journal of Islamic Religious Education*, 2(1), 57-76. <https://doi.org/10.30762/ed.v2i1.582>
- Fathani, A.H. (2019). Pembelajaran matematika bagi santri pondok pesantren berbasis kecerdasan majemuk. *ANARGYA : Jurnal Ilmiah Pendidikan Matematika*, 2(1), 48–55. <https://doi.org/10.24176/anargya.v2i1.3043>
- Gembong, S., Krisdiana, I., Suprpto, E., Setyansah, R.K., Widodo, S.A., & Murtafiah, W. (2022). The effectiveness of mathematics learning based on javanese vegetable salad context to Improving students' higher order thinking skills. *Pegegog Journal of Education and Instruction*, 13(1), 84–91. <https://doi.org/10.47750/pegegog.13.01.10>
- Hanafi, Y., Taufiq, A., Saefi, M., Ikhsan, M.A., Diyana, T.N., Thoriquttyas, T., & Anam, F.K. (2021). The new identity of indonesian islamic boarding schools in the "new normal" : the education leadership response to covid-19. *Heliyon*, 7(3), 1–10. <https://doi.org/10.1016/j.heliyon.2021.e06549>
- Heong, Y.M., Othman, W.B., Yunos, J.B.M., Kiong, T.T., Hassan, R., & Mohamad, M.M. (2012). The need analysis of learning higher order thinking skills for generating ideas. *Procedia - Social and Behavioral Sciences*, 59, 195-203. <https://doi.org/10.1016/j.sbspro.2012.09.265>
- Ismail, R., Retnawari, H., Sugiman, Arovah, N.I., & Imawan, O.R. (2024). Contexts proposed by teachers in Papua for developing mathematics HOTS assessment instruments: a phenomenological study. *Journal of Education and E-Learning Research*, 11(3), 548–556. <https://doi.org/10.20448/jeelr.v11i3.5922>
- Karwadi, K., Zakaria, A.R., & Syafii, A. (2024). A review of the effects of active learning on high order thinking skills : a meta-analysis within islamic education. *Journal of Education and Learning (EduLearn)*, 18(1), 97–106. <https://doi.org/10.11591/edulearn.v18i1.20895>
- Kunanti, E.S. (2020). Penyusunan pengembangan penilaian berbasis HOTS. *Prosiding Seminar Nasional PBSI-III Tahun 2020*, 19–26. <https://digilib.unimed.ac.id/id/eprint/41215/>
- Manik, P., Saraswati, S., Ngurah, G., & Agustika, S. (2020). Kemampuan berpikir tingkat tinggi dalam menyelesaikan soal HOTS mata pelajaran matematika. *Jurnal Ilmiah Sekolah Dasar*, 4(2), 257–269. <https://doi.org/10.23887/jisd.v4i2.25336>
- Milenia, D., Resti, N.C., & Rahayu, D.S. (2022). Kemampuan siswa SMP dalam penyelesaian soal matematika berbasis HOTS pada materi pola bilangan. *Jurnal Ilmiah Matematika Realistik*, 3(2), 100–108. <https://doi.org/10.33365/ji-mr.v3i2.2297>
- Nenohai, J.M.H., Samo, D.D., Garak, S.S., Rimo, I.H.E., & Soko, I.P. (2024). Preservice mathematics teacher knowledge of higher order thinking skills. *Journal of Education and Learning (EduLearn)*, 18(2), 588–597. <https://doi.org/10.11591/edulearn.v18i2.21184>
- Nowlan, N., Arya, A., Qorbani, S., & Abdinejad, M. (2032). Higher-order thinking skills assessment in 3d virtual learning environments using motifs and expert data. *Computers & Education X Reality*, 2(5), 100012. <https://doi.org/10.1016/j.cexr.2023.100012>
- Pilten, G. (2010). Evaluation of the skills of 5 th grade primary school students ' high-order

- thinking levels in reading. *Procedia : Social and Behavioral Sciences*, 2(2), 1326–1331. <https://doi.org/10.1016/j.sbspro.2010.03.195>
- Purbaningrum, K.A. (2017). Kemampuan berpikir tingkat tinggi siswa smp dalam pemecahan masalah matematika ditinjau dari gaya belajar. *JPPM*, 10(2), 40–49. <https://doi.org/10.30870/jppm.v10i2.2029>
- Rahmi, U., & Azrul. (2022). Optimizing the discussion methods in blended learning to improve student's high order thinking skills. *Pegem Journal of Education and Instruction*, 12(3), 190–196. <https://doi.org/10.47750/pegegog.12.03.20>
- Ramdhani, S., Suryadi, D., & Prabawanto, S. (2021). Hambatan belajar matematika di pondok pesantren. *Jurnal Analisa*, 7(1), 46–55. <https://doi.org/10.15575/ja.v7i1.10106>
- Sugiyono. (2013). *Metode penelitian pendidikan pendekatan kuantitatif, kualitatif, dan r&d*. Bandung: Alfabeta.
- Thanheiser, E. (2023). What is the mathematics in mathematics education? *Journal of Mathematical Behavior*, 70(1), 1–13. <https://doi.org/10.1016/j.jmathb.2023.101033>
- Vijayaratnam, P. (2012). Developing higher order thinking skills and team commitment via group problem solving : a bridge to the real world. *Procedia - Social and Behavioral Sciences*, 66(1), 53–63. <https://doi.org/10.1016/j.sbspro.2012.11.247>
- Yaniawati, P., Maat, S.M., Supianti, I.I., & Fisher, D. (2021). Mathematics mobile blended learning development: student-oriented high order thinking skill learning. *European Journal of Educational Research*, 11(1), 69–81. <https://doi.org/10.12973/eu-jer.11.1.69>