INTERNATIONAL JOURNAL OF GEOMETRY RESEARCH AND INVENTIONS IN EDUCATION
Vol. 01 No. 02 (2025), 1-14

GRADIENT

https://journals.eduped.org/index.php/gradient
E-ISSN 3036-959X

Designing a Geometry Examination Framework to Evaluate Higher-Order Thinking Skills

Nia Kania[1*], Iik Nurhikmayati[2], Vahid Noouzi Larsari[3]

[1,2] Universitas Majalengka, Indonesia
[3]Charles University, Czech Republic

*Corresponding author: niakania@unma.ac.id

Article Info

Received December 19, 2024
Revised January 14, 2025
Accepted February 17, 2025

Abstract

This study addresses the growing need to enhance students' higher-order thinking skills (HOTS) in mathematics, a critical aspect for fostering problem-solving and critical reasoning abilities essential for academic success. As the demand for more advanced cognitive skills increases, assessing and strengthening HOTS in education becomes crucial. This study aims to develop a valid, practical, and effective instrument for assessing higher-order thinking skills (HOTS) in junior high school mathematics. Employing a research and development (R&D) approach, the study adapts Mardapi's (2008) development model, which includes nine stages. However, this research focuses on the first seven stages, leaving the final implementation to school teachers. The instrument was tested on 40 students at one of the junior high schools in Rajagaluh, Majalengka Regency. The validity analysis demonstrated that all V values exceeded 0.3, indicating strong content validity. Reliability testing yielded a Cronbach's Alpha score of 0.52, suggesting moderate internal consistency. The findings revealed that students' HOTS performance remains suboptimal, highlighting the need for targeted instructional strategies. Strengthening HOTS can be achieved through structured practice with complex problem-solving tasks. This study provides a robust framework for educators to assess and enhance students' critical thinking abilities in mathematics.

Keywords: Geometry Test; Higher-order Thinking Skills, Mathematics Assessment, Instrument Development, Junior High School, HOTS Measurement

How to cite: Kania, N., Nurhikmayati, I., and Larsari, V. N. (2025). Designing a Geometry Examination Framework to Evaluate Higher-Order Thinking Skills. *Journal of Geometry Research and Innovation in Education*, 02(1), 1-14, doi. https://doi.org/10.56855/gradient.v2i02.1327

1. Introduction

The representation of geometric data is notoriously challenging due to its requirement for an in-depth understanding of spatial reasoning and shapes. The vast diversity of shapes, sizes, and locations involved in geometry makes it common practice to perceive objects in three dimensions. This is essential because of the intrinsic nature of the subject matter. Furthermore, the challenge is exacerbated by linking abstract concepts with concrete visual representations. For instance, mental imagery becomes a crucial tool when studying the relationships between geometric shapes or attempting to grasp transformations such as rotation and reflection. Consequently, visual aids such as graphic software, physical models, and diagrams are frequently employed to help students understand geometric concepts by simplifying ideas that would otherwise be difficult to comprehend.

Higher-order Thinking Skills (HOTS) have become a national priority, especially in mathematics education. The development of HOTS in students is vital for their success in mathematics learning in the classroom (Gavronskaya et al., 2022; Kania & Kusumah, 2025; Sadijah et al., 2021). This is in line with the Minister of Education Regulation No. 22 of 2006, which refers to the Content Standards and emphasizes the significance of mathematics in cultivating students' logical, analytical, systematic, critical, and creative thinking abilities, as well as their capacity for communication. Among the key HOTS capabilities are application, analysis, synthesis, and evaluation (Tambunan & Naibaho, 2019). The implementation of HOTS in Indonesian classrooms follows Bloom's taxonomy, with the cognitive order including Applying, Analyzing, Evaluating, and Creating (Bertucio, 2017; Chandio et al., 2021; Radmehr & Drake, 2019; Samo et al., 2020; Setiawan et al., 2021; Spindler, 2020; Sujoko & Darmawan, 2013).

In response to the 2011 TIMSS and 2012 PISA reports, Indonesia's Ministry of Education introduced the 2013 curriculum (Kurniawan et al., 2021). This curriculum was designed to enhance students' HOTS (Haniah et al., 2020). Strengthening HOTS in education is essential, as fostering these skills is a goal outlined in the Basic Framework and Curriculum Structure for Junior High Schools (SMP/MTs) (Indriyana & Kuswandono, 2019; Rahmi et al., 2021; Setiawan et al., 2021; Tambunan & Naibaho, 2019; Tyas et al., 2020). Moreover, the competency standards for primary and secondary education graduates (Ministry of Education and Culture, 2022) highlight the importance of these skills.

To develop HOTS effectively, students must engage in activities that foster these skills. Higher-order reasoning typically demands more cognitive processing processing (Mitani, 2021; Samo & Kartasasmita, 2017) compared to other types of thinking (Kim, 2020). HOTS is a process in which individuals receive new knowledge and then utilize it to solve problems (Mulyatna et al., 2021). The nature of HOTS suggests that these are mental tasks that are non-routine, more complex, and require substantial effort (Yaniawati et al., 2021). Research by Putranta et al., (2021) emphasizes that developing HOTS requires consistent and repeated practice. Therefore, it is crucial to equip students with the tools to tackle complex problem-solving scenarios and continuously evaluate their thinking processes (Ansari et al., 2021). These skills are inherently linked to the multifaceted nature of HOTS, which often involves multiple solutions.

In the 2013/2014 National Examination (UN), the government introduced questions designed to assess HOTS. Approximately 20% of the National Examination questions now include HOTS-related content (Haniah et al., 2020). This inclusion reflects the centrality of HOTS in the 2013 curriculum (Harta et al., 2020). The composition of questions includes 10%-15% focusing on reasoning (Heryani et al., 2023), 50%-60% on application, and 25%-30% on knowledge and understanding (Kemendikbud, 2019). These evaluation tasks are specifically designed to assess students' mastery of HOTS (Friyatmi et al., 2020). According to Pi'i, (2016) teachers are expected to incorporate HOTS questions to evaluate students' academic achievements and meet the intellectual demands of the curriculum. Therefore, teachers must be capable of developing valid and reliable HOTS instruments based on these principles.

The National Examination, as mentioned above, serves as one of the instruments teachers can refer to when assessing HOTS. This highlights the need for teachers to develop or adapt HOTS instruments, which measure various aspects of students' higher-order thinking. The quality of teachers today directly influences the future quality of education (Kania et al., 2020). Teachers must demonstrate creativity in crafting questions that challenge students to think critically and engage in higher-level thinking (Syafryadin et al., 2021). The teacher's role in developing HOTS is fundamental in both planning and implementation (Bayram, 2021; Simsek, 2021). This should be done consistently and with careful consideration (Ramdiah et al., 2019). The goal is to assess students' relative strengths and weaknesses in higher-order thinking (Collins, 2014).

Effective teaching requires a deep understanding of problem-solving, the application of concepts, and the ability to analyze and evaluate information produced through communication, reasoning, observation, reflection, and experience (Rahman et al., 2021). Teachers can also use HOTS assessments to gauge students' preparedness for the National Examination. If teachers neglect this aspect, the potential development of students' HOTS may remain unrecognized and undeveloped.

Given the importance of this matter, the researcher argues that developing HOTS instruments is essential. Numerous studies have demonstrated the effectiveness of learning methods in enhancing students' HOTS, highlighting the need for robust HOTS assessment tools. There are two key reasons why HOTS is crucial: first, it enables students to succeed academically; and second, it prepares them to become adults who contribute positively to society (Conklin, 2012). Schoenfeld, (2010) asserts that HOTS is fundamental for everyone and should be taught to help students think creatively, critically, and skeptically about issues in the 21st century. In an era of competitive knowledge acquisition, students need to process and store information, connect it to prior knowledge, and analyze it to solve real-world problems (Abdurrahman, 2021; Aulya et al., 2020).

The researcher hopes that by developing this HOTS instrument, it will assist teachers who may not be familiar with how to create or adapt HOTS assessments. This will, in turn, encourage teachers to consistently incorporate HOTS questions into their teaching, enabling them to evaluate students' HOTS abilities effectively. Operationally, this study aims to: (1) produce HOTS indicators for mathematics at the junior high school level.

2. Methods

This research includes research and development (R&D), where researchers create a measuring instrument for students' higher order thinking skills (HOTS). This research yielded a product in the form of an instrument that the teacher can use as an example or reference to measure students' HOTS in mathematics. This research model was adapted from Mardapi, (2008) development model, which consists of the following steps: (1) gathering test specifications, (2) writing test questions, (3) studying test questions, (4) conducting test trials, (5) analyzing the items, (6) improving the test, (7) assembling the test, (8) carrying out the test, and (9) interpreting the test results. In this study, however, the researcher restricted the development model to the steps of assembling the test. In other words, the researcher did not conduct the test and did not interpret the results of the test. The researchers have high hopes that teachers who are also educators in schools will be able to carry out the final two steps of the process. This instrument was tested on a group of 40 students at at one of the junior high schools in Rajagaluh, Majalengka Regency.

The developed instrument focuses on measuring students' HOTS in mathematics, targeting indicators across various cognitive levels, including application, analysis, evaluation, and creation, based on the principles of Bloom's Taxonomy. The instrument is divided into categories based on HOTS abilities, which can be detailed in the table below:

Table 1 - Targeting indicators across various cognitive levels

Indicator	Cognitive level	Description	Number of items
Application	Applying	Applying mathematical concepts to solve real-world problems.	1,2
Analysis	Analyzing	Breaking down complex problems into simpler components.	3
Evaluation	Evaluating	Critically evaluating different methods to solve problems.	4
Creation	Creating	Designing original mathematical solutions or models.	5

The data analysis technique used in this study is content validity, construct, reliability, difficulty level, discrimination, and students' mathematical HOTS abilities. The content validity technique asks an expert, in this case as a validator, to examine the accuracy and provide an assessment of the suitability of the item and its indicators, and the editor of the question composition.

Following the expert's assessment, the researcher calculated the outcomes of the assessment using a validity index, incorporating the Aiken index, as follows:

$$V = \frac{\sum s}{N(c-1)}, \qquad s = r - l$$

(Aiken, 1980, p.956).

Information:
Possible V range of numbers
r;appraiser rating
l: lowest category rater rating
c: highest category
N: number of raters/respondents

The range of numbers is 0 to 1. The greater the validity of an item, the higher the number V near 1 or equal to 1; conversely, the greater the validity of an item, the lower the number V near 0 or equal to 0. Similarly, the number of items on the market has decreased (Aiken, 1980)

In addition, exploratory factor analysis was used to prove construct validity. The percentage of variance seen from the Kaiser Meyer Olkin (KMO) result indicates exploratory factor analysis (Taştan & Yilmaz, 2008). The IBM SPSS 20 Software was used to calculate the KMO value. In addition, the sample that was taken makes it possible to conduct additional research (Santoso, 2006).

The Cronbach's alpha formula and IBM SPSS 20 were used to estimate the instrument's reliability for an internal consistency test. Cronbach's Alpha values of 0.60 or less 1 indicate that the instrument has met the reliable criteria, whereas a value of less than 0.50 indicates the instrument is not reliable (Basuki & Hariyanto, 2014; Surapranata, 2009). Meanwhile, the differentiating power of description questions is calculated using the formula:

$$DB = \frac{top\ group\ mean - bottom\ group\ mean}{maximum\ score\ of\ the\ question}$$

(Nitko & Brookhart, 2011)

After the numbers have been crunched, they are separated into three piles: accepted, revised, and rejected. (Table 1). It depends on the coefficient of difference. If there are questions that are rejected, they can be discarded or replaced with new items.

Table 2 - Differential power coefficient

Differential Power Coefficient	**Category**
DB > 0,3	Received
0,10 ≤ DB < 0,30	Revised
DB < 0,10	Rejected

(Surapranata, 2009)

After the instrument met the criteria set out above, data analysis was conducted to determine the students' HOTS. The criteria used are if the assessment results show the student's mathematics HOTS score is more than or equal to 65 (on a scale of 0-100), then the student's mathematics HOTS has met the good criteria.

3. Results and Discussion

3.1. Results

3.1.1. *Development of Test Specifications*

At this stage, the researcher designed the specifications for an instrument to measure students' Higher-order Thinking Skills (HOTS) in mathematics. The instrument was developed based on cognitive indicators from Bloom's Taxonomy, which include: application, analysis, evaluation, and creation.

3.1.2. *Item Writing*

Five items were developed based on the specifications and HOTS indicators. Each item was designed to reflect a specific cognitive level from Bloom's Taxonomy.

3.1.3. *Item Review*

Two mathematics experts reviewed the instrument developed as validators. The initial review indicated that the instrument required revisions. Revisions were made based on the reviewers' feedback, and the instrument was then resubmitted for final evaluation.

3.1.4. *Content Validity Testing*

Two experts in mathematics examined the instrument, resulting in proof of the instrument's content validity. The study results showed that the initial instrument made by the researcher was not good. Therefore, the researcher improved/revised as much as possible according to the suggestions written by the validator on the instrument sheets.

After completion of the repair, the instrument is returned to the validator who evaluates each item. After revisions, content validity was assessed using the Aiken index. The Aiken values for each item were as follows:

Table 3 – Aiken values for each item

Item Number	Aiken Value	Description
1	0.625	Valid
2	0.75	Valid
3	0.5	Sufficiently Valid
4	0.875	Highly Valid
5	0.625	Valid

All items were considered valid as the Aiken values exceeded 0.5. Thus, they were deemed suitable for field testing.

3.1.5. *Instrument Trial*

The instrument was tested on 40 students from a junior high school in Rajagaluh District, Majalengka Regency. The trial results were analyzed for construct validity, reliability, item difficulty, and discriminatory power.

3.1.6. Construct Validity

Construct validity was tested through exploratory factor analysis (EFA). The results of the EFA revealed that the instrument explained 56.9% of the variance, indicating that the items effectively represented the measured HOTS construct.

3.1.7. Reliability Testing

Reliability was assessed using Cronbach's Alpha, yielding a coefficient of 0.52, which indicates adequate reliability, though it is not yet optimal. This suggests that the internal consistency of the items is sufficient for exploratory research purposes.

3.1.8. Item Difficulty Analysis

The difficulty level of each item was analyzed to ensure variation in the questions. The results are as follows:

Table 4 – Difficulty index

Item Number	Difficulty Index	Category
1	0.34	Moderate
2	0.45	Moderate
3	0.34	Moderate
4	0.53	Moderate
5	0.13	Difficult

Most of the items fell within the moderate category, with one item classified as difficult, indicating a heterogeneous nature of the instrument.

3.1.9. *Discriminatory Power Analysis*

Discriminatory power indicates how well the items distinguish between high- and low-performing students. The analysis yielded the following results:

Table 5 – Discriminatory power of items

Item Number	Discriminatory Power	Category
1	0.25	Revise
2	0.21	Revise
3	0.27	Revise
4	0.34	Accept
5	0.20	Revise

Based on these results, no items were rejected, but several items required revision to enhance their discriminatory power.

3.1.10. Final Instrument Development

After undergoing analysis, the instrument was revised based on empirical data. The five items were refined with minor adjustments to wording and context based on the feedback from the analysis and validators. The instrument is now considered ready for use by teachers as a tool to assess students' HOTS in mathematics.

3.2. Discussions

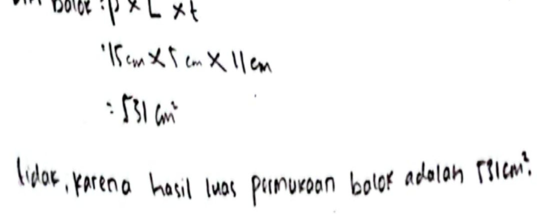
3.2.1. Student Math HOTS

In general, the analysis results of all student scores indicate that the higher-order thinking (HOT) mathematical ability of students in the trial is still below 65, with an average score of 26.38 on a scale of 100. This average shows that the students in the trial were not very good at higher-order thinking (HOT) in math. Students are not used to answering HOTS questions, so they do not know what to do. Everyone's ability to think is different, so everyone needs to practice and improve their math skills (A'yun et al., 2021).

However, HOTS is a government priority that students must have in mathematics. However, students do not get enough practice on HOTS questions in everyday learning. So that students' skills are only at the level of formal mathematics.

Meanwhile, the characteristics of HOTS require strong analytical skills in solving mathematical problems. In learning activities, students must be taught to be able to think critically, highly, and on their own (Asari et al., 2019).

The following is an illustration of student answers:



Explanation:
Is known:
Cuboid =Length × Width × Height
=15 cm x 5 cm x 11 cm
No, because the result of the surface area of The block is 521 cm^2

Figure 1 Student answer

In Figure 1, the student's answer illustrates that the student does not know the formula for surface area. Although the answer "no" is the correct answer, the reason students answer "no" is the wrong answer. So that students can not give proper reasons.

Figure 1 represents the level of analytical thinking and is the lowest level of thinking based on HOTS Bloom's.

The results of the analysis of the student's mathematical HOTS, as described above, also illustrate that the student's HOTS is still low. This means that the things students have done to learn so far have not been geared toward helping them build their HOTS. This aligns with what researchers have found by Riadi & Retnawati (2014), each of whom develops teaching materials and learning tools oriented to students' HOTS. The HOTS ability of students before participating in the HOTS-oriented learning (pretest) is still very low (Kania, Suryadi, et al., 2024; Musfiqi & Jailani, 2014).

The next research stage is instrument improvement. Instrument repair is carried out based on input from the Expert. In addition, the improvement of the instrument is also based on the stages of analysis that have been carried out, such as the level of difficulty and distinguishing power. Researchers divided the instruments tested into three categories. The first category is the accepted instrument category, the second category is the revised instrument category, and the third category is the rejected instrument category.

Based on the results of the Expert validation regarding advanced validity, there are several improvements related to the instrument grid. Meanwhile, for content validity, the expert validator assessed that the instrument was able to measure what was intended to be measured. So that improvements are only to the instrument grid and language structure of the questions. The following are comments from expert validators:

Table 7 - Expert validator suggestions

No Question	Face Validity Comment	Comments Content Validity
1	The indicators are insufficiently specific, such that if they were created by someone else, the problem could take on a different form in the space in question.	Questions can measure the indicator's stages of ability.
2	The name of the space must be included in the indicator so that others can ask comparable questions.	
3, 4, 5	The building's name must be displayed on the indicator.	

Following the correction of the revised items, the next step is to assemble/compile the items. to form what is known as the HOTS instrument. At this stage, the researcher grouped the questions into a single instrument. The following are examples of questions on the developed instrument:

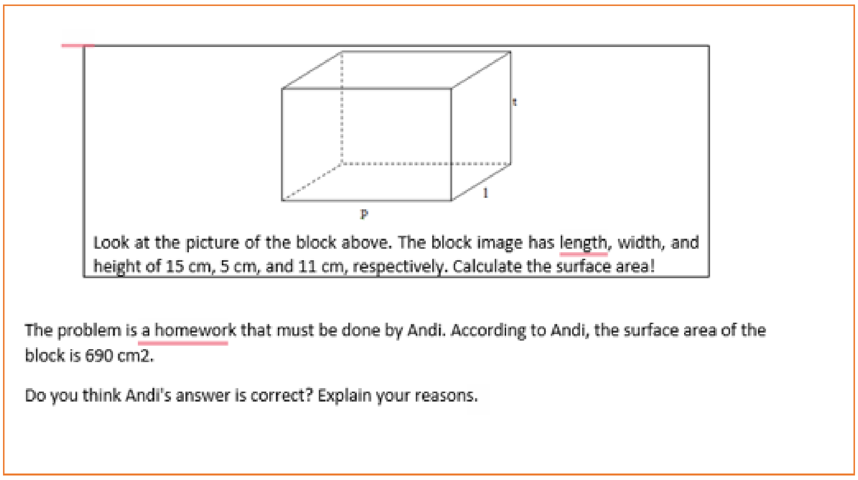


Figure 2 Question

The question grid is corrected according to the expert's advice. The indicator grid is improved from the language side, which only creates one perception. So that anyone who reads the grid can make the expected questions well.

In supporting students' difficulties in solving HOTS questions, teachers are also required to be able to develop HOTS measuring instruments properly. A teacher is not only tasked with conveying information, but also requires the ability to design effective learning to achieve learning objectives. Cooperation should also be developed in the field of secondary education. The implementation of HOTS places a significant amount of responsibility and obligation on educators (Jasmina et al., 2022; Maryani et al., 2022). Through the thought process, a person is invited to construct the knowledge he has (Ratnawulan & Kania, 2020). Material mastery, high strategy, and technical evaluation are all very necessary requirements, and they must be considered in terms of affective, cognitive, and psychomotor dimensions (Arifin & Bonyah, 2024; Asari et al., 2019; Kania & Kusumah, 2025). Activities in education include the delivery of subject matter, such as providing explanations to students, and instructions for each student. Other fundamental things needs that need to be considered, such as guidance and direction, also need to be considered. In conclusion, we are of the firm belief that the principles of teaching organization that are practiced by educators in each and every school ought to be given top priority. This would ensure the improvement of the quality of students, teachers, educational institutions, and the education system (Kullan et al., 2022). This is because education involves more than just the delivery of subject matter (A'yun et al., 2021; Kania, Kusumah, et al., 2024).

4. Conclusions

Based on the researchers' analysis findings, it is possible to conclude that the higher-order thinking skills (HOTS) measuring instrument for junior high school mathematics is valid, as all V values are greater than 0.3. The instrument, which consists of five items, was put through its paces. The instrument was divided into two test packages. The item packages produced by test packages A and B were reliable, with Cronbach's Alpha scores of 0.52. Furthermore, the results of the instrument trial indicate that higher-order thinking skills

(HOTS) in mathematics for class VIII students are poor. The average value of the test results that are less than 65, which is 26.38 on a scale of 100, indicates this. This is due to students' decreased ability to answer HOTS questions. To improve HOTS, students are expected to be able to practice questions that are high-level types of questions.

Acknowledgments
Thank you to Universitas Majalengka for the grants for this research.

Conflict of Interest
The authors declare no conflicts of interest.

References
A'yun, Q., Yanuarti, R. S., & Cahyo, D. A. (2021). Performance Analysis of Students Through Critical Thinking Ability Based on Mathematic Ability. *Malikussaleh Journal of Mathematics Learning (MJML)*, 4(1), 23–27. https://doi.org/10.4108/eai.19-12-2020.2309168

Abdurrahman, M. S. (2021). Improving polytechnic students' high-order-thinking-skills through inquiry-based learning in mathematics classroom. *International Journal of Evaluation and Research in Education*, *10*(3), 976–983. https://doi.org/10.11591/IJERE.V10I3.21771

Aiken, L. R. (1980). *Content validity and reliability of single items or questionnaires*. Educational and Psychological Measurement.

Ansari, B. I., Saleh, M., Nurhaidah, & Taufiq. (2021). Exploring students' learning strategies and self-regulated learning in solving mathematical higher-order thinking problems. *European Journal of Educational Research*, *10*(2), 743–756. https://doi.org/10.12973/eu-jer.10.2.743

Arifin, Z., & Bonyah, E. (2024). Tracing the Roots of Error: A Polya Method Analysis on Student Problem Solving in Curved Surface Solids. *International Journal of Geometry Research and Inventions in Education (Gradient)*, *1*(01), 21–30. https://doi.org/10.56855/gradient.v1i01.1144

Asari, S., Husniah, R., Ma'rifah, U., & Anwar, K. (2019). Fostering Students' High Order Thinking Skills through the Use of Interpretation Cards. *International Journal of Education and Literacy Studies*, *7*(4), 17. https://doi.org/10.7575/aiac.ijels.v.7n.4p.17

Aulya, N. R., Ichsan, I. Z., & Rahman, M. (2020). HOTS Analysis to Develop E-Supplement Book Based on Plant Physiology Research. *Universal Journal of Educational Research*, *8*(12B), 8461–8466. https://doi.org/10.13189/ujer.2020.082654

Basuki, I., & Hariyanto. (2014). *Asesemen pembelajaran*. PT. Remaja Rosdakarya.

Bayram, S. (2021). World journal on educational technology: Current issues. *World Journal on Educational Technology: Current Issues*, *13*(4), 82–95. https://doi.org/https://doi.org/10.18844/wjet.v13i1.5401

Bertucio, B. (2017). The Cartesian Heritage of Bloom's Taxonomy. *Studies in Philosophy and Education*, *36*(4), 477–497. https://doi.org/10.1007/s11217-017-9575-2

Chandio, M. T., Nishat Zafar, & Solangi, G. M. (2021). Bloom's Taxonomy: Reforming Pedagogy Through Assessment. *Journal of Education and Educational Development*, *8*(1), 109–140. https://doi.org/10.22555/joeed.v8i1.308

Collins, R. (2014). Skills for the 21st Century: teaching higher-order thinking. *Curriculum & Leadership Journal*, *12*(14).

http://www.curriculum.edu.au/leader/teaching_higher_order_thinking,37431.html?issueID=12910

Conklin, W. (2012). *Strategies for Developing Higher-Order Thinking Skills, Grades 6-12: Grades 6-12*. Shell Educational Publishing.Inc.

Friyatmi, F., Mardapi, D., & Haryanto, H. (2020). Assessing Students' Higher Order Thinking Skills Using Multidimensional Item Response Theory. *Problems of Education in the 21st Century*, *78*(2), 196–214. https://doi.org/10.33225/pec/20.78.196

Gavronskaya, Y., Larchenkova, L., Berestova, A., Latysheva, V., & Smirnov, S. (2022). The Development of Critical Thinking Skills in Mobile Learning: Fact-Checking and Getting Rid of Cognitive Distortions. *International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE)*, *10*(2), 51–68. https://doi.org/10.23947/2334-8496-2022-10-2-51-68

Haniah, A. R., Aman., & Setiawan, R. (2020). Integration of strengthening of character education and higher order thinking skills in history learning. *Journal of Education and Learning (EduLearn)*, *14*(2), 183–190. https://doi.org/10.11591/edulearn.v14i2.15010

Harta, J., Rasuh, N. T., & Seriang, A. (2020). Using HOTS-Based Chemistry National Exam Questions to Map the Analytical Abilities of Senior High School Students. *Journal of Science Learning*, *3*(3), 143–148. https://doi.org/10.17509/jsl.v3i3.22387

Heryani, Y., Kartono, Wijayanti, K., & Dewi, N. R. (2023). Analysis of Student's Mathematical Literacy Ability in Solving HOTS Problems in Minimum Competency Assessment. *Journal of Higher Education Theory and Practice*, *23*(16). https://doi.org/10.33423/jhetp.v23i16.6470

Indriyana, B. S., & Kuswandono, P. (2019). Developing Students Higher Order Thinking Skills (HOTS) in Reading: English Teachers Strategies in Selected Junior High Schools. *JET (Journal of English Teaching)*, *5*(3), 204. https://doi.org/10.33541/jet.v5i3.1313

Jasmina, A., Alla, B., Yulia, T., Grosseck, G., & Živkov, A. M. (2022). The Quality of Online Higher Education Teaching During the Covid-19 Pandemic. *International Journal of Cognitive Research in Science, Engineering and Education*, *10*(1), 47–55. https://doi.org/10.23947/2334-8496-2022-10-1-47-55

Kania, N., & Kusumah, Y. S. (2025). THE MEASUREMENT OF HIGHER-ORDER THINKING SKILLS: A SYSTEMATIC LITERATURE REVIEW. *Malaysian Journal of Learning and Instruction*, *22*(1), 97–116. https://doi.org/10.32890/mjli2025.22.1.6

Kania, N., Kusumah, Y. S., Dahlan, J. A., Nurlaelah, E., & Kyaruzi, F. (2024). Decoding Student Struggles in Geometry: Newman Error Analysis of Higher-Order Thinking Skills. *International Journal of Geometry Research and Inventions in Education (Gradient)*, *1*(01), 31–47. https://doi.org/10.56855/gradient.v1i01.1146

Kania, N., Nurhikmayati, I., & Suciawati, V. (2020). Pre-service mathematics teachers' experiences of teaching practice in function composition. *Journal of Physics: Conference Series*, *1613*(1). https://doi.org/10.1088/1742-6596/1613/1/012013

Kania, N., Suryadi, D., Kusumah, Y. S., Dahlan, J. A., Nurlaelah, E., & Elsayed, E. E. (2024). Comparative Praxeology: Assessing High-Level Cognitive Skills in TIMSS and Indonesian National Examinations. *International Journal of Applied Learning and Research in Algebra*, *1*(1), 25–47. https://doi.org/10.56855/algebra.v1i1.1160

Kemendikbud. (2019). *Tingkat Kesulitan Soal UN 2019 Tidak Berubah, Ini Komposisi Soalnya*. https://www.kemdikbud.go.id

Kim, H. J. (2020). Students' academic use of mobile technology and higher-order thinking skills: The role of active engagement. *Education Sciences*, *10*(3).

https://doi.org/10.3390/educsci10030047
Kullan, S., Mansor, M., & Ishak, R. (2022). The validity and reliability of an instrument to evaluate the practices of learning organization. *International Journal of Evaluation and Research in Education (IJERE)*, *11*(4), 1725. https://doi.org/10.11591/ijere.v11i4.22974
Kurniawan, H., Budiyono, Sajidan, & Siswandari. (2021). The PINTER learning model to enhance higher order thinking and communication skill in Algebra. *International Journal of Instruction*, *14*(3), 359–374. https://doi.org/10.29333/iji.2021.14321a
Mardapi, D. (2008). *Technique of preparation of test and non test instrument.* Mitra Cendekia Press.
Maryani, I., Prasetyo, Z. K., Wilujeng, I., & Purwanti, S. (2022). Promoting higher-order thinking skills during online learning: The integration of metacognition in science for higher education. *International Journal of Evaluation and Research in Education (IJERE)*, *11*(4), 1980. https://doi.org/10.11591/ijere.v11i4.23129
Mitani, H. (2021). Test Score Gaps in Higher Order Thinking Skills : Exploring Instructional Practices to Improve the Skills and Narrow the Gaps. *AERA Open*, *7*(1). https://doi.org/10.1177/23328584211016470
Mulyatna, F., Imswatama, A., & Rahmawati, N. D. (2021). Design Ethnic-Math HOTS: Mathematics Higher Order Thinking Skill Questions Based On Culture and Local Wisdom. *Malikussaleh Journal of Mathematics Learning (MJML)*, *4*(1), 48. https://doi.org/10.29103/mjml.v4i1.3059
Musfiqi, A., & Jailani. (2014). Pengembangan Bahan Ajar Matematika yang Berorientasi pada Karakter dan Higher Order Thinking Skill (HOTS) Developing Mathematics Instructional Materials Oriented to Character and Higher Order Thinking Skill (Hots). *PYTHAGORAS: Jurnal Pendidikan Matematika*, *9*(1), 45–59. http://journal.uny.ac.id/index.php/pythagoras
Nitko, A. J., & Brookhart, S. M. (2011). *Educational assessment of students* (Xth editio). Prentice Hall Englewood Cliffs.
Pi'i. (2016). Developing High-Level Thinking Learning and Assessment in High School History Subjects (in bahasa),". *Jurnal Sejarah Dan Budaya,*.
Putranta, H., Supahar, Setiyatna, H., Choiriyah, S., Dwandaru, W. S. B., Widodo, & Warsono. (2021). The Effect of Smartphone Usage Intensity on High School Students' Higher Order Thinking Skills in Physics Learning. *Journal of Turkish Science Education*, *18*(3), 421–438. https://doi.org/10.36681/tused.2021.82
Radmehr, F., & Drake, M. (2019). Revised Bloom's taxonomy and major theories and frameworks that influence the teaching, learning, and assessment of mathematics: a comparison. *International Journal of Mathematical Education in Science and Technology*, *50*(6), 895–920. https://doi.org/10.1080/0020739X.2018.1549336
Rahman, N. A., Rosli, R., Rambely, A. S., & Halim, L. (2021). Mathematics teachers' practices of stem education: A systematic literature review. *European Journal of Educational Research*, *10*(3), 1541–1559. https://doi.org/10.12973/EU-JER.10.3.1541
Rahmi, Y. L., Habibah, I. N., Zulyusri, Z., & Darussyamsu, R. (2021). HOTS assessment in circulatory system learning: Validity, reliability, and item quality. *JPBI (Jurnal Pendidikan Biologi Indonesia)*, *7*(2), 171–178. https://doi.org/10.22219/jpbi.v7i2.15513
Ramdiah, S., Abidinsyah, Royani, M., & Husamah. (2019). Understanding, planning, and implementation of HOTS by senior high school biology teachers in Banjarmasin-Indonesia. *International Journal of Instruction*, *12*(1), 425–440. https://doi.org/10.29333/iji.2019.12128a

Ratnawulan, N., & Kania, N. (2020). Implementation of Cooperative Learning Model Numbered Head Together (Nht) Type To Improve Learning Activities. *Jurnal THEOREMS (The Original Research of Mathematics)*, *4*(2), 161–168.

Riadi, A., & Retnawati, H. (2014). Pengembangan Perangkat Pembelajaran untuk Meningkatkan HOTS pada Kompetensi Bangun Ruang Sisi Datar Developing Learning Kit to Improve HOTS for Flat Side of Space Competence. *PYTHAGORAS: Jurnal Pendidikan Matematika*, *9*(2), 126–135. http://journal.uny.ac.id/index.php/pythagoras

Sadijah, C., Murtafiah, W., Anwar, L., Nurhakiki, R., & Cahyowati, E. T. D. (2021). Teaching higher-order thinking skills in mathematics classrooms: Gender differences. *Journal on Mathematics Education*, *12*(1), 159–179. https://doi.org/10.22342/jme.12.1.13087.159-180

Samo, D. D., Garak, S. S., & Maifa, T. S. (2020). Mathematics teacher knowledge in higher-order thinking skill: curriculum, pedagogy, and assessment. In *Journal of Physics: Conference Series* (Vol. 1663, p. 12012). IOP Publishing. https://doi.org/10.1088/1742-6596/1663/1/012012

Samo, D. D., & Kartasasmita, B. (2017). *Developing Contextual Mathematical Thinking Learning Model to Enhance Higher-Order Thinking Ability for Middle School Students*. *10*(12), 17–29. https://doi.org/10.5539/ies.v10n12p17

Santoso, S. (2006). *Seri Solusi Bisnis Berbasis TI: Menggunakan SPSS untuk Statistik Multivariat*. Elex Media Komputindo.

Schoenfeld, A. (2010). *How We Think*. University of California-Berkeley.

Setiawan, J., Sudrajat, A. A., Aman, A., & Kumalasari, D. (2021). Development of higher order thinking skill assessment instruments in learning Indonesian history. In *International Journal of Evaluation and Research in Education (IJERE)* (Vol. 10, Issue 2, p. 545). Institute of Advanced Engineering and Science. https://doi.org/10.11591/ijere.v10i2.20796

Simsek, Z. Z. (2021). "Is It Valid or Not?": Pre-Service Teachers Judge the Validity of Mathematical Statements and Student Arguments. *European Journal of Science and Mathematics Education*, *9*(2), 26–42. https://doi.org/10.30935/SCIMATH/10772

Spindler, R. (2020). Aligning Modeling Projects with Bloom's Taxonomy. *Primus*, *30*(5), 601–616. https://doi.org/10.1080/10511970.2019.1619208

Sujoko, E., & Darmawan, I. P. A. (2013). REVISI TAKSONOMI PEMBELAJARAN BENYAMIN S. BLOOM I Putu Ayub Darmawan. *Jurnal Satya Widya*, *29*(1), 30–39.

Surapranata, S. (2009). *Analisis, validitas, reliabilitas, dan interpretasihasil tes implementasi kurikulum 2004.* PT. Remaja Rosdakarya.

Syafryadin, Harahap, A., Haryani, & Astrid, A. (2021). Boosting classroom interaction based on higher order thinking skills (Hots) in english learning for beginners. *International Journal of Language Education*, 5(1), 477–489. https://doi.org/10.26858/IJOLE.V5I1.15211

Tambunan, H., & Naibaho, T. (2019). Performance of mathematics teachers to build students' high order thinking skills (HOTS). *Journal of Education and Learning (EduLearn)*, *13*(1), 111–117. https://doi.org/10.11591/edulearn.v13i1.11218

Taştan, M., & Yilmaz, K. (2008). Organizational citizenship and organizational justice scales' adaptation to turkish. *Egitim ve Bilim*, *33*(150), 87–96.

Tyas, M. A., Nurkamto, J., & Marmanto, S. (2020). Cultivating Students ' Higher -Order Thinking Skills in EFL Classes : The Role of The Teacher and The Textbook. *International Online Journal of Education and Teaching (IOJET)*, *7*(1), 267–276. http://iojet.org/index.php/IOJET/article/view/773

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/https://doi.org/10.12973/eu-jer.11.1.69