

DEVELOPMENT OF HIGHER-ORDER THINKING SKILL (HOTS) TEST ON MATHEMATICS IN SECONDARY SCHOOL

Ega Gradini¹, Cut Khairunnisak², Julia Noviani^{3*}

^{1,3*} IAIN Takengon, Takengon, Indonesia

² Universitas Syiah Kuala, Banda Aceh, Indonesia

*Corresponding author

E-mail: ega.gradini@iaintakengon.ac.id¹⁾

cut.khairunnisak@unsyiah.ac.id²⁾

julianoviani@iaintakengon.ac.id^{3*)}

Received 27 December 2021; Received in revised form 15 March 2022; Accepted 22 March 2022

Abstract

Assessing students' higher-order thinking skills are still challenging for mathematics teachers in Indonesia. This paper aims to develop a valid, reliable, practical, and effective test to assess students' higher-order thinking skills in mathematics learning. The Plomp's developmental model, Generic Model for Educational Design, was deployed to develop the test. The method consisted of 4 phases; (1) Problem analysis, (2) Design, (3) Realization, and (4) Evaluation and Implementation. The quality of the developed test was examined by using Plomp's product quality criteria; (1) Validity, (2) Reliability, (3) Practicality/usability, and (4) Effectiveness. The instruments in this study were validation sheets and practical response questionnaires. The test was validated by expert validation involving two experts on HOTS in mathematics and a trial involving 35 grade 7 secondary school students. The practicality/usability of the test was examined by 25 mathematics teachers from different secondary schools. The effectiveness was measured by trial on three secondary schools. The validity and reliability were measured using Gregory's Expert Agreement Index (EAI). The rubric's practicality was analyzed using practicality product criteria, and the effectivity of the test was analyzed using a two-tailed *t*-test. The research shows that the validity of the test is 0.88 (high validity), and the reliability is 0.93 (reliable). The practicality of the test is 75.33 (practical). The trial on the lower-index, middle-index, and higher-index schools revealed that the test effectively assessed students' higher-order thinking skills. In conclusion, the HOTS test is valid, reliable, practical, and effective to measure students' higher-order thinking skills in mathematics learning.

Keywords: Assessment, HOTS, HOTS problems, thinking skills.

Abstrak

Menilai keterampilan berpikir tingkat tinggi siswa masih menjadi tantangan bagi guru matematika di Indonesia. Penelitian ini bertujuan untuk mengembangkan tes yang valid, reliabel, praktis, dan efektif untuk menilai kemampuan berpikir tingkat tinggi siswa dalam pembelajaran matematika. Model pengembangan Plomp, the Generic Model for Educational Design, digunakan untuk mengembangkan tes. Metode Plomp terdiri dari 4 fase; (1) Analisis masalah, (2) Perancangan, (3) Realisasi, dan (4) Evaluasi dan Implementasi. Kualitas produk yang dikembangkan diperiksa dengan menggunakan kriteria kualitas produk dari Plomp; (1) Validitas, (2) Reliabilitas, (3) Kepraktisan/kegunaan, dan (4) Efektivitas. Instrumen pada penelitian ini adalah lembar validasi dan angket respon praktikalitas. Produk tes HOTS divalidasi melalui; (1) validasi ahli yang melibatkan dua validator ahli HOTS matematika dan (2) uji coba yang melibatkan 35 siswa SMP kelas VIII. Praktikalitas tes diuji oleh 25 guru matematika dari SMP yang berbeda. Efektivitas diukur dengan uji coba di tiga SMP/MTs. Validitas dan reliabilitas diukur dengan menggunakan Gregory's Expert Agreement Index (EAI). Kepraktisan rubrik dianalisis menggunakan kriteria produk kepraktisan, dan keefektifan produk dianalisis menggunakan uji-t two-tailed. Hasil penelitian menunjukkan bahwa validitas tes adalah 0,88 (validitas tinggi), dan reliabilitas 0,93 (reliable). Kepraktisan tes mencapai 75,33 (Praktis). Uji coba di sekolah ber-indeks kompetensi rendah, menengah, dan tinggi menemukan bahwa tes HOTS yang dikembangkan efektif untuk mengukur HOTS siswa. Kesimpulannya, tes HOTS yang dikembangkan pada penelitian ini valid, reliabel, praktis, dan efektif dalam mengukur kemampuan berpikir tingkat tinggi siswa dalam pembelajaran matematika.

Kata kunci: Asesmen, HOTS, keterampilan berpikir, soal HOTS.



DOI: <https://doi.org/10.24127/ajpm.v11i1.4649>

INTRODUCTION

The paradigm of education in the 21st Century has evolved from achieving basic literacy abilities (reading, writing, and arithmetic) to emphasizing soft skills. As the global reform, the assessment of learning achievements is also evolving, encouraging a change from conventional instruction for algorithmic, from lower-order thinking skills (LOTs) toward higher-order thinking skills (Barak & Dori, 2009). The 2013 curriculum, as explained in Permendikbud No. 35 of 2018, was developed to answer various challenges of globalization development. The curriculum expects students to develop 1) critical thinking and problem-solving skills, 2) communication and collaboration skills, 3) creativity and innovation skills, 4) information and communication technology literacy, 5) contextual learning skills, and 6) information and media literacy skills (Rosdiana et al., 2020). Besides, the curriculum also emphasized Higher Order Thinking Skills (Gradini et al., 2018). Indonesia's participation in international assessment, such as the Program for International Student Assessment (PISA), is also part of the challenges of globalization in curriculum development (Izzati et al., 2020). However, many schools still have difficulty implementing the 2013 curriculum, particularly in mathematics (Gradini, 2019; Sari et al., 2019).

Even though the curriculum was developed based on a thorough evaluation of TIMSS and PISA results (Fitriati et al., 2021), Indonesian students' mathematical literacy score is consistently lower than the average international standard score (OECD, 2019; Stacey, 2011), and even decreased in PISA 2018. Based on the

PISA 2018 result (OECD, 2019), about 28% of Indonesian students only reached Level 2 out of six levels of mathematical literacy. Only about 1% of Indonesian students were found to have HOTS and reached Level 5 or above, whereas the average of OECD was 11%. It means that only 1% of Indonesian students can model complex situations mathematically and select, compare, and evaluate appropriate problem-solving strategies to deal with them. The result of PISA 2018 implies that the students did not meet the expectations of Indonesia's current mathematics curriculum. It means that the students could interpret and identify, without direct instruction, how simple situations can be represented mathematically (e.g., comparing total distances across two alternative routes or converting prices into different currencies).

The low skill level in mathematical literacy may be caused by teachers' lack of knowledge of such skills (Tanujaya, 2016). Retnawati (Retnawati et al., 2018) discovered that mathematics teachers in Indonesian were still unfamiliar with the concept of HOT. Brookhart (Brookhart, 2010) and Tanujaya & Doorman (Tanudjaya & Doorman, 2020) define the skill itself as the ability to transfer knowledge into new contexts, apply critical and creative thinking, and problem-solving. As a transfer of knowledge, HOTS is defined as thinking skills in the cognitive domains of C4 (analysis), C5 (evaluate), and C6 (create) from Bloom's Taxonomy (Anderson et al., 2001). As critical and creative thinking, Brookhart suggested that higher-order thinking can be achieved by fostering critical thinking. Some of the skills that become indicators of critical thinking are comparing, differentiating, estimating,

DOI: <https://doi.org/10.24127/ajpm.v11i1.4649>

drawing conclusions, influencing, generalizing, specializing, classifying, grouping, sorting, predicting, validating, proving, connecting, analyzing, evaluating, and making patterns. Meanwhile, HOTS as a problem-solving process makes students solve real problems in real life, which are generally unique so that the completion procedures are also unique and not routine (Brookhart, 2010; Brookhart & Nitko, 2011).

Students' development of HOTS can be achieved through teachers' active role in planning, implementing, and evaluating HOTS-based learning (Retnawati et al., 2018). However, some previous studies indicated that teachers faced obstacles and difficulties while training their students about HOTS. The obstacles encountered by mathematics teachers include designing HOTS-based problems and finding suitable learning instruments (Jaelani & Retnawati, 2016; Retnawati et al., 2017), and assessing students' HOTS (Retnawati et al., 2016). Furthermore, school exams typically assess students' HOTS after students' LOTs, whereas teachers should analyze students' HOTS independently from their LOT (Jansen & Möller, 2022).

Given the difficulties teachers had in assessing and developing students' HOTS, the availability of HOTS-based mathematics learning resources and instruments is essential and number of researchers have reported their work on developing HOTS-based test instruments (Faridah et al., 2018; Kurniasi & Arsisari, 2020; Masniladevi et al., 2019; Sadiada et al., 2018; Susetyawati & Nuryani, 2021; Tanujaya, 2016). However, despite numerous studies on developing HOTS-based mathematics learning tools in modules, student worksheets, and

evaluation instruments, the number of HOTS-based mathematics learning resources available is still insufficient. The developed products were also inadequate for covering the entire scope of subjects taught in junior high school, i.e., numbers, algebra, geometry and measurement, and statistics. Furthermore, HOTS-based mathematics learning materials for numbers, especially the topic of Patterns and Number Sequences, are limited.

Because there is still a gap between the ideal situation and the empirical situation in the field, particularly in numbers, it is necessary to develop qualified HOTS-based assessment instruments. Therefore, this study intends to develop a valid, reliable, practical, and effective test for assessing students' higher-order thinking skills in mathematics learning.

METHOD

The Plomps' developmental model, Generic Model for Educational Design, was deployed to develop the test. The method consisted of 4 phases; (1) Problem analysis, (2) Design, (3) Realization, and (4) Evaluation and Implementation (Tjeerd Plomp, 2000). The problem analysis was conducted through 5 stages: problem analysis, students' condition analysis, learning material analysis, task analysis, and specification of learning objectives. In the design phase, the prototype of HOTS is designed by generating all the parts of the solution, comparing and evaluating the various alternatives then producing the best design choice of the test. The test was designed and developed to assess students' higher-order thinking skills in secondary mathematics learning. In the realization phase, the test was constructed using the HOTS aspect defined by Brookhart, namely the

DOI: <https://doi.org/10.24127/ajpm.v11i1.4649>

top three levels of Bloom's taxonomy; logical reasoning; and problem-solving (Brookhart, 2010). HOTS-based questions require students' ability to reason and solve mathematical problems that are tested as a tool to measure students' mathematical literacy skills (Dinni, 2018). Using HOTS category questions aims to reveal students' creative and critical abilities by developing detailed knowledge obtained from multiple sources while deciding on a problem (Izzati et al., 2020).

The developed test's quality was examined using Plomps' product quality criteria; validity, reliability, practicality/usability, and effectiveness (Nieveen & Folmer, 2013). Two expert validators validated the test on HOTS in mathematics. The trial involved 35 students from Aceh Tengah's secondary schools. Further, the practicality/usability of the test was examined by 25 mathematics teachers from 8 secondary schools in Aceh Tengah, Aceh. Then the effectiveness was measured by trial on three secondary schools that represented the higher, middle, and lower competency schools. The trial was carried out in 3 secondary schools. Then numerous students were selected from the upper, middle, and lower competency schools; 86, 34, 68, respectively, using random sampling. The validity and reliability were measured using Gregory's Expert Agreement Index (EAI) (Gregory, 2011). The practicality was analyzed using practicality product criteria, and the test's effectivity was analyzed using a two-tailed *t-test*.

The test consisted of 5 HOTS problems on Number Patterns and Sequence taught to 8th grader students. The test assesses the following cognitive level; analyze, evaluate, and

create. Test the difficulty level of the question using item analysis. Boateng defines difficulty level as the proportion of students who properly answer a question. It has a range of 0.0 to 1.0. A high difficulty score indicates that a larger percentage of the sample successfully answered the question. A lower difficulty score indicates that a smaller percentage of the sample successfully answered the question (Boateng et al., 2018).

RESULTS AND DISCUSSION

1. The Problem Analysis Phase

The first phase of the Plomp Generic model for Educational Design is problem analysis. A Focused Group Discussion (FGD) with six mathematics teachers was organized to investigate teachers' obstacles in constructing HOTS problems. Teachers found that constructing and solving the HOTS problem is still challenging due to their lack of knowledge and training. Teachers are also not accustomed to solving non-routine and higher cognitive problems. Another obstacle is to elevate the Indicators of Competency Achievement from lower order to higher one.

2. The Design Phase

The second phase is design. In this phase, relevant concepts were identified and compiled systematically. Then, the test blueprint was designed by developing the questions grid based on the Basic Competencies that students need to achieve.

3. The Realization Phase

The third phase is the realization phase. In this phase, the HOTS test is constructed by integrating HOTS characteristics as suggested by Brookhart. Recall that according to Brookhart, assessing HOTS means assessing HOTS as a transfer of

DOI: <https://doi.org/10.24127/ajpm.v11i1.4649>

knowledge, critical thinking, and problem-solving (Brookhart, 2010; Brookhart & Nitko, 2011). This test consisted of 5 problems on C4

(analysis), C5 (evaluate), and C6 (create) as follows. All these questions required critical thinking and problem-solving skills.

Table 1. The higher-order thinking skills test

Problem No	Level of Cognitive	Characteristic
#1 The Dots #2 The Relationship between x and y	C4 - Analyze	<ul style="list-style-type: none"> Break knowledge/material or concept into parts and determine how the parts relate to each other or the overall structure or purpose. Do mental actions include distinguishing, organizing, and attributing between components or parts on this cognitive level. Illustrating the mental function by creating a spreadsheet, survey, chart, diagram, or graphical representation. Identify and relate relevant data/information from the situation/problem and make the correct conclusions from the information given
#3 Folding Paper #4 The Bouncing Ball	C5 - Evaluate	<ul style="list-style-type: none"> Make judgments based on criteria and standards through checking and criticism. Criticisms, recommendations, and reports are some of the products that can be made to demonstrate the evaluation process. Assess solutions, ideas, and methodologies using suitable criteria or existing standards to ensure their effectiveness or benefits. Unite elements into a coherent or functional whole; reorganization of elements into new patterns or structures through generation, planning, or production.
#5 Tetrahedra	C6 - Create	<ul style="list-style-type: none"> Putting the parts together in a new way or synthesizing the parts into something new and different to create a new form or product.

4. Evaluation and Implementation Phase

The last phase is implementation and evaluation that conducted in a cycle. In other words, if the test does not meet the requirement of the product quality (validity, reliability, practicality, and effectivity), the constructed test is revised, implemented, and evaluated until it meets the requirement. The following results from the implementation and evaluation phase of the HOTS test's validity, reliability, practicality, and effectivity.

A. Validity and Reliability

Aspects validated in the HOTS test are its relevance to material, language, time allocation, and HOTS characteristics. The results of expert validation on the HOTS test are shown in the Table 2.

Table 2. Expert Validation of HOTS Test

Experts Agreement	Validator I	
	Weak Relevance	Strong Relevance
Weak Relevance	0	1
Strong Relevance	1	15

The expert agreement on validation results shows that the HOTS test validity coefficient is 0.88. It indicates that the test has high validity. Then, the trial on a limited group was conducted to find its reliability, level of difficulty, and discriminating power of the HOTS problems. The trial was conducted on 35 secondary school students in Aceh Tengah. The results of the trial are described in Table 3.

Table 3. Validity and Reability of HOTS Test

DOI: <https://doi.org/10.24127/ajpm.v11i1.4649>

	Value
Mean	35.309
Standard Deviation	19.117
Median	30.000
Reliability Coefficient	0.930
Standard of Error	5.838
Average of Difficulty Level	0.368
Average of Discriminating Power	0.483

The reliability coefficient of the HOTS test is 0.930, and the standard of error for the test is 5.838. It is indicated that the HOTS test has high reliability. The average difficulty level of the questions is 0.368 that indicates the questions are on a moderate level. The average discriminating power is 0.483, and the questions on this test are on a good level. In other words, the test can distinguish students from the upper and lower academic achievement groups.

B. Practicality

Table 4. The Practicality of HOTS Test

Aspect	Scale	
1 Language	4,12	Very Good
2 Structure and layout	3,68	Good
3 Instruction	3,96	Good
4 Relevance to Material	3,96	Good
5 Time Allocation	4,16	Good
6 HOTS characteristics	3,92	Good

A 5-scale questionnaire has been distributed to 25 Mathematics teachers of 8 secondary schools in Aceh Tengah. The findings show that the practicality of the HOTS test is on a 'very practical' level, with a score of practicality 79.33. The practicality measures language, structure and layout, instruction, relevance to material, time allocation, and HOTS characteristics. Around 84%

of teachers agree that the language in the HOTS test is easy to understand. 56% of teachers found that the layout makes it easier to read and understand. However, 12% of teachers suggest that the problems should merge with the description.

The questionnaire also revealed that 64% of teachers consider the instructions for doing the HOTS test presented clearly, and 72% of teachers state that the test follows the material being studied by students. Further, 72% of teachers state that time allotted is sufficient to solve all the problems. Lastly, 68% of teachers stated that all problems assess levels C4 (analysis), C5 (evaluate), and C6 (create) of Bloom's taxonomy correctly and adequately.

C. Effectivity

The trial was conducted in 3 secondary schools. Then numerous students were selected from the upper, middle, and lower competency schools; 86, 34, 68, respectively, using random sampling. The result of the measurement is as follows.

The trial on small groups conducted at three secondary schools shows that the HOTS test effectively measures students' higher-order thinking skills in upper, middle, and lower competency schools. Thus, the test is applicable to measure students in Aceh Tengah's secondary schools. However, larger groups are needed for future research.

Table 5. Effectiveness of HOTS Test

School Competency	Effectivity Test			HOTS Level			
	t	df	Sig. (2-tailed)	Exemplary	Proficient	Develop	Emerging
Upper	8.072	68	.010	17,65%	23,53%	11,76%	47,06%
Middle	3.424	36	.002	15,79%	21,05%	52,63%	10,53%
Lower	4.358	84	.023	4,65%	16,28%	18,60%	60,47%

DOI: <https://doi.org/10.24127/ajpm.v11i1.4649>

The trial on upper, middle and lower competency schools showed that the t-value of upper, middle and lower competency schools is 8.072, 3.424, and 4.358, respectively. Meanwhile, the value of Sig. (2-tailed) is 0.010, 0.002, and 0.023 respectively. These values are less than the predefined significance value, 0.05. Thus, the test significantly effective measures students' higher-order thinking skills.

The test also showed the students' HOTS levels, namely 1) Exemplary, 2) Proficient, 3) Develop, and 4) Emerging. Students in the Exemplary, Proficient, and Develop stages are categorized as students at a high HOTS level, while students in the emerging stage are still at a low one. The test showed that in lower competency schools, most students' higher-order thinking skills are in the emerging level. This finding follows (Hartini et al., 2021; Yuliati & Lestari, 2018) who found that most students had low HOTS levels. This finding is also in line with Kurniati's finding that argues that the high-order thinking skills of junior high school students are at medium and low levels in solving HOTS Mathematics questions (Kurniati et al., 2016). This finding is also supported by (Hartatiana et al., 2020) who found that most students (73%) were at a low HOTS level, and (Nafi'an & Pradani, 2019), who found not more than 50% of students were able to solve HOTS-type mathematical problems.

The upper competency schools have more students at the exemplary level than middle and lower ones. This finding indicates a significant difference between the higher-order thinking skills of students from upper, middle, and low competency schools. This finding follows (Hajar et al., 2018; Maharaj & Wagh, 2016; Tasman, 2020) who found

that students' HOTS ability was influenced by school classification. Students from upper classification schools tend to answer HOTS questions better than students from the lower one. Furthermore, this finding is also in line with (Budsankom et al., 2015) finding that the school and classroom environment significantly affect students' HOTS abilities. These findings indicate that the developed HOTS test is valid, practical, and effective for use in Mathematics classrooms. Teachers are encouraged to utilize this test to assess students' higher-order thinking skills in the Pattern and Number Sequence material taught on 8th grade.

Problem #1. The Dots

Notice the following pattern

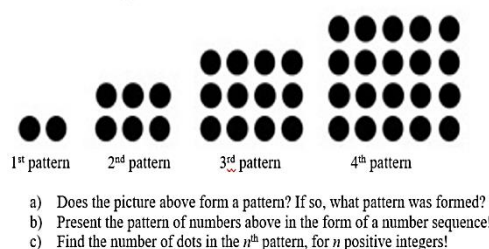


Figure 1. HOTS problem #1 on C4 level

This problem assessed students' ability in making generalizations from patterns in number sequences and object configuration sequences. It requires students to analyze the formation of patterns in a sequence of object configurations. The problem is on the C4 (analyze) level. Students have to identify and relate relevant data/information from the situation/ problem, make the correct conclusions from a set of data/information, assess the quality/accuracy of a statement or argument, and detect consistency and inconsistency in a process/ product with evidence.

The information given on this problem is the dots that form a pattern. Students were asked to analyze the

DOI: <https://doi.org/10.24127/ajpm.v11i1.4649>

pattern formed and determine it using their knowledge. Students were asked to present the pattern in a number sequence and find the number of dots in the n th pattern for n positive integers. This problem has a validity of 0.84 (high). Boateng refers to difficulty level as the percentage of students who answered a question correctly (Boateng et al., 2018). On average, the difficulty level of problem #1 is moderate. The difficulty level of question #a is 0.91 (easy), #b is 0.68 (moderate), and #c is 0.51 (moderate).

Problem #2. The Relationship between x and y .

The following table explains the relationship between 2 numbers.

1	1
2	...
3	...
4	7
5
6
7	13
x	y

- What is the correct number to fill in ... in the table?
- Can you explain the relationship between x and y ?

Figure 2. HOTS problem #2 on C4 level

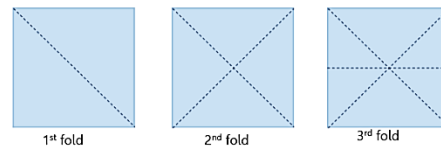
This problem assessed students' ability to make generalizations from patterns in number sequences and object configuration sequences. It requires students to analyze and find patterns in a number sequence. The problem is on the C4 (analyze) level that requires students to identify and relate relevant data/information from the situation/problem and make the correct conclusions from the information given.

Students are given the relationship between 2 numbers on the table on this problem. They asked to fill in the missing number presented in the n th term of the unknown number sequence. Then, they were asked to explain the relationship between x and y using written mathematical expressions. They need to provide evidence to support their

justification to answer this problem. The validity of problem #2 is 0.92 (high). The level of difficulty of question #a is 0.78 (moderate) and #b is 0.82 (easy).

Problem #3. Folding Paper

Take the paper given by the teacher and fold it according to the following folds. The dotted line is the fold line. Notice the number of triangles that are formed!



- Fill in the following table with the number of triangles formed in each fold. You can ignore the size of each triangle.

Fold	The Number of Triangles
1 st fold	2
2 nd fold	4
3 rd fold	6
4 th fold	...
5 th fold	...
6 th fold	...
7 th fold	...

- What pattern is formed?
- Make a sequences of numbers formed from the pattern!
- Find the 10th and 100th terms of the number sequence!

Figure 3. HOTS problem #3 on C5 level

The problem is on the C5 level that assesses students' ability to solve contextual problems using number patterns. Students must assess solutions, ideas, and methodologies to solve the Folding Paper problem. They have to use suitable criteria or existing standards to ensure their evidence supports the ideas. By folding the paper, students provided the term of sequence and formed the pattern. Consequently, they can find the particular term of the sequence. The validity of this problem is 0.89 (high). The level of difficulty of question #a is 0.91 (easy), #b is 0.75 (moderate), #c is 0.72 (moderate), and #d is 0.23 (difficult).

Problem #4. The Bouncing Ball.

(Adapted from

<https://mathigon.org/course/sequences/a-rithmetic-geometric>)

DOI: <https://doi.org/10.24127/ajpm.v11i1.4649>

- Barak, M., & Dori, Y. J. (2009). Enhancing higher order thinking skills among inservice science teachers via embedded assessment. *Journal of Science Teacher Education*, 20(5), 459–474.
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in Public Health*, 6, 149.
- Brookhart, S. M. (2010). How To Assess Higher-Order Thinking Skills in Your Classroom. In *Association for Supervision and Curriculum Development*. Association for Supervision and Curriculum Development. <https://doi.org/10.1177/002205741808801819>
- Brookhart, S. M., & Nitko, A. J. (2011). Strategies for Constructing Assessments of Higher-Order Thinking Skills. In G. Schraw & D. H. Robinson (Eds.), *Assessment of Higher Order Thinking Skills* (pp. 153–168). Information Age and Publishing. <https://doi.org/10.4018/978-1-7998-2314-8.ch008>
- Budsankom, P., Sawangboon, T., Damrongpanit, S., & Chuensirimongkol, J. (2015). Factors affecting higher order thinking skills of students: A meta-analytic structural equation modeling study. *Educational Research and Reviews*, 10(19), 2639–2652.
- Dinni, H. N. (2018). HOTS (High Order Thinking Skills) dan Kaitannya dengan Kemampuan Literasi Matematika. *Prisma*, 1, 170–176.
- Faridah, R., Siswono, T. Y. E., & Rahaju, E. B. (2018). Developing Higher Order Thinking Skill (HOTS) Mathematic Problem Using That Quiz Application. *Mathematics, Informatics, Science, and Education International Conference (MISEIC 2018)*, 167–170.
- Fitriati, F., Marlaini, M., & Elizar, E. (2021). Integrating Rich Task into the Mathematics Classroom to Develop Students ' Higher Order Thinking Skills : A Collaborative Action Research Study in a Secondary School. *Ilkogretim Online - Elementary Education Online*, 20(1), 479–494. <https://doi.org/10.17051/ilkonline.2021.01.042>
- Gradini, E. (2019). Menilik konsep kemampuan berpikir tingkat tinggi (higher order thinking skills) dalam pembelajaran matematika. *Jurnal Numeracy*, 6(2), 189–203.
- Gradini, E., Firmansyah B, & Noviani, J. (2018). Menakar Kemampuan Berpikir Tingkat Tinggi Calon Guru Matematika Melalui Level HOTS Marzano. *Eduma: Mathematics Teaching and Learning*, 7(2), 41–48.
- Gregory, R. J. (2011). *Psychological Testing: History, Principles, and Applications* (7th editio). Pearson.
- Hajar, Y., Yanwar, R., Jalaludin, M. A., Achmad, N., Indriani, G. S., Hidayat, W., & Rohaeti, E. E. (2018). Analisis Kemampuan High Order Thinking (Hot) Siswa Smp Negeri Di Kota Cimahi. *JPMI (Jurnal Pembelajaran Matematika Inovatif)*, 1(3), 453–458.
- Hartatiana, Wardani, A. K., & Megawati, M. (2020). Kemampuan Berpikir Tingkat Tinggi Siswa SMP dalam Menyelesaikan Soal Matematika Model PISA. *Jurnal*

DOI: <https://doi.org/10.24127/ajpm.v11i1.4649>

- Pendidikan Matematika*, 14(01), 15–24.
- Hartini, P., Setiadi, H., & Ernawati, E. (2021). Instrumen penilaian berbasis LOTS dan HOTS buatan guru kelas VI sekolah dasar mata pelajaran IPA di Jakarta. *Jurnal Penelitian Dan Penilaian Pendidikan*, 3(1), 14–24.
- Izzati, N., Antika, R., Asma, N., & Siregar, R. (2020). Pembimbingan Guru Dalam Mengembangkan Soal. *JMM Jurnal Masyarakat Mandiri*, 4(3), 370–381. <https://doi.org/https://doi.org/10.31764/jmm.v4i3.2511>
- Jaelani, J., & Retnawati, H. (2016). The Challenges of Junior High School Mathematic Teachers in Implementing the Problem-Based Learning for Improving the Higher-Order Thinking Skills. *Online Journal of Counseling & Education*, 5(3).
- Jansen, T., & Möller, J. (2022). Teacher judgments in school exams: Influences of students' lower-order-thinking skills on the assessment of students' higher-order-thinking skills. *Teaching and Teacher Education*, 111, 103616.
- Kurniasi, E. R., & Arsisari, A. (2020). Pengembangan Instrumen Pengukur Higher Order Thinking Skills (HOTS) Matematika Pada Siswa Sekolah Menengah Pertama. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, 9(4), 1213–1222.
- Kurniati, D., Harimukti, R., & Jamil, N. A. (2016). Kemampuan berpikir tingkat tinggi siswa SMP di Kabupaten Jember dalam menyelesaikan soal berstandar PISA. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 20(2), 142–155.
- Maharaj, A., & Wagh, V. (2016). Formulating tasks to develop HOTS for first-year calculus based on Brookhart abilities. *South African Journal of Science*, 112(11–12), 1–6.
- Masniladevi, M., Oneika, S. E., & Helsa, Y. (2019). Development of HOTS Test Instruments as an Effort to Improve the Reasoning Ability of Elementary School Students by Using Edugames. *5th International Conference on Education and Technology (ICET 2019)*, 682–686.
- Nafi'an, M. I., & Pradani, S. L. (2019). Analisis Kemampuan Pemecahan Masalah Siswa dalam Menyelesaikan Soal Matematika Tipe Higher Order Thinking Skill (HOTS). *Kreano, Jurnal Matematika Kreatif-Inovatif*, 10(2), 112–118.
- Nieveen, N., & Folmer, E. (2013). Formative Evaluation in Educational Design Research. In T Plomp & N. Nieveen (Eds.), *Netherlands Institute for Curriculum Development: SLO* (pp. 152–169). Netherlands Institute for Curriculum Development: SLO.
- OECD. (2019). Programme for International Student Assessment (PISA) results from PISA 2018. In *Indonesia Country Note - PISA 2018 Results*.
- Plomp, Tjeerd. (2000). Development Research in / on Educational Development. *CITE Seminar Series*, 04.
- Retnawati, H., Djidu, H., Kartianom, A., & Anazifa, R. D. (2018). Teachers' knowledge about higher-order thinking skills and its learning strategy. *Problems of Education in the 21st Century*,

DOI: <https://doi.org/10.24127/ajpm.v11i1.4649>

- 76(2), 215.
- Retnawati, H., Hadi, S., & Nugraha, A. C. (2016). Vocational High School Teachers' Difficulties in Implementing the Assessment in Curriculum 2013 in Yogyakarta Province of Indonesia. *International Journal of Instruction*, 9(1), 33–48. <https://doi.org/10.12973/iji.2016.914a>
- Retnawati, H., Munadi, S., Arlinwibowo, J., Wulandari, N. F., & Sulistyaningsih, E. (2017). Teachers' difficulties in implementing thematic teaching and learning in elementary schools. *New Educational Review*, 48(2), 201–212. <https://doi.org/10.15804/ner.2017.48.2.16>
- Rosdiana, M., Sumarni, S., & Siswanto, B. (2020). Implementation of 21st Century Learning Through Lesson Study. *4th International Conference on Arts Language and Culture (ICALC 2019)*, 346–353.
- Sadieda, L. U., Indayati, T., Faizien, M. I., & others. (2018). Developing an assessment instrument of higher order thinking skills in mathematics with in Islamic context. *Journal of Physics: Conference Series*, 1097(1), 12151.
- Sari, Y., Cahyaningtyas, A. P., Maharani, M. M., Yustiana, S., & Kusumadewi, R. F. (2019). Meningkatkan kemampuan menyusun soal IPA berorientasi HOTS bagi guru Sekolah Dasar Gugus Pandanaran Dabin IV UPTD Semarang Tengah. *Indonesian Journal of Community Services*, 1(2), 175. <https://doi.org/10.30659/ijocs.1.2.175-183>
- Stacey, K. (2011). The PISA View of Mathematical Literacy in Indonesia. *Journal of Research on Mathematics Education*, 2(2), 95–126.
- Susetyawati, M. M. E., & Nuryani, C. E. (2021). Developing Mathematics HOTS Test Items in Essay. *Journal of Physics: Conference Series*, 1823(1), 12041.
- Tanudjaya, C. P., & Doorman, M. (2020). EXAMINING HIGHER ORDER THINKING IN INDONESIAN LOWER. *Journal on Mathematics Education*, 11(2), 277–300.
- Tanujaya, B. (2016). Development of an Instrument to Measure Higher Order Thinking Skills in Senior High School Mathematics Instruction. *Journal of Education and Practice*, 7(21), 144–148.
- Tasman, F. (2020). Improving Junior High School Teacher Ability in Designing Higher order Thinking (HOTS) Problems. *Pelita Eksakta*, 3(1), 25–30.
- Yuliati, S. R., & Lestari, I. (2018). Higher-order thinking skills (hots) analysis of students in solving hots question in higher education. *Perspektif Ilmu Pendidikan*, 32(2), 181–188.