



THE DEVELOPMENT OF HOTS (HIGHER-ORDER THINKING SKILLS) ASSESSMENT INSTRUMENT BASED ON NUMERATION LITERACY IN MATHEMATICS FOR JUNIOR HIGH SCHOOL

Nurul Wardani^{1*}, Tanti Jumaisyaroh Siregar²

^{1,2}Department of Mathematics Education, Faculty of Tarbiyah and Teacher Training, Universitas Islam Negeri Sumatera Utara, Indonesia

*Corresponding author: nurulwardani0@gmail.com

Article Info

Article history:

Received: June 14, 2023

Accepted: July 29, 2023

Published: July 31, 2023

Keywords:

Assessment instrument
 Higher-order thinking skills
 HOTS
 Numeration literacy

ABSTRACT

This research aims to develop a HOTS (Higher-Order Thinking Skills) assessment instrument based on numeracy literacy. This research employed a research and development method with 4D models. Feasibility analysis, discriminating index, item difficulty level, and product practicality were used in data collection techniques. The results showed that the average value of the item validity was 85.1% which indicated that this product is suitable for use as an assessment instrument. The discriminating index test obtained a value of 0.11 in the weak category, and the item difficulty level test obtained a value of 60% in the moderate category. Furthermore, The product practicality test on 30 students yielded an average score of 67.25%, which means that this product is practical. Thus, the instrument developed is feasible and practical as a stimulus to train higher-order thinking skills and increase cognitive knowledge in mathematics learning. Future research is expected to develop instruments in other mathematical materials.

PENGEMBANGAN INSTRUMEN ASESMEN HOTS (HIGHER ORDER THINKING SKILLS) BERBASIS LITERASI NUMERASI PADA MATA PELAJARAN MATEMATIKA SMP

ABSTRAK

Kata Kunci:

Instrumen asesmen
 Kemampuan berpikir tingkat tinggi
 HOTS
 Literasi numerasi

Tujuan penelitian ini ialah mengembangkan instrumen asesmen HOTS (*Higher Order Thinking Skills*) berbasis literasi numerasi. Penelitian ini menggunakan metode *research and development* dengan model 4D. Analisis kelayakan, daya pembeda, tingkat kesukaran butir soal dan kepraktisan produk digunakan dalam teknik pengumpulan data. Hasil penelitian menunjukkan nilai rata-rata validitas butir soal sebesar 85,1% yang menunjukkan bahwa produk ini layak digunakan sebagai instrumen penilaian. Berdasarkan hasil uji daya pembeda sebesar 0,11 dalam kategori lemah dan hasil uji tingkat kesukaran butir soal sebesar 60% dalam kategori sedang. Sementara berdasarkan hasil uji kepraktisan produk 30 siswa memberi nilai rata-rata 67,25% yang berarti produk ini praktis digunakan. Dengan demikian, instrumen yang dikembangkan layak dan praktis digunakan sebagai stimulus untuk melatih kemampuan berpikir tingkat tinggi serta dapat meningkatkan pengetahuan kognitif pada pembelajaran matematika peserta didik. Penelitian selanjutnya diharapkan dapat mengembangkan instrument pada materi matematika lainnya.

1. INTRODUCTION

The industrial revolution 4.0 is the era of intelligence and the progress of science and technology. Mankind has entered the era of civilization with strict competition. Progress became one of the supporting factors of the revolution of the industrial revolution 4.0 [1]. This progress makes science a supporting factor for innovation related to science and technology. Mathematics as part of education cannot be separated from science and technology, which also plays a role in progress.

Mathematics plays a key role in educational development and progress, and its application can be used in various fields. It differs from other subjects because it is characteristic [2]. A key characteristic of mathematics is a deductive thought pattern, a thought pattern based on a chronological order of concepts, axioms (postulates), definitions, theorems, propositions, and applications in mathematics, other areas, and everyday life. Such traits are applied to special things [3]. Based on these characteristics, mathematics is understanding concepts and arithmetic and the competence to reason and analyze.

The competence to reason and analyze can be achieved by continually assessing. This assessment tests for lower cognitive and material mastery [4]. From the perspective of collection and processing, assessments profoundly affect education, especially in class, because, through the internal assessment, the educators observe learners by their ability. The assessment of the class is prepared to dig up knowledge about students' study activities and experiences [5]. The usual assessments in the class are tests that measure the cognitive aspects of learners, including multiple-choice tests and descriptions that measure learners' higher-order thinking skills [6].

Based on the above opinion, the researchers surveyed one of the MTs in Medan in the eighth grade of the 2022/2023 academic year. The researchers found some issues that cause students not to have HOTS, such as a lack of concept explanation with HOTS abilities in everyday life. Teachers or educators also often depend on guidebooks and learning methods that do not match the current needs of students. As a result, the students are not familiar with completing the problem that focuses on HOTS problems related to daily life. According to the researcher's observation results, most students did not have higher-order thinking skills. Concretional samples can be seen from learners who have not had their HOTS. They have difficulties when answering HOTS-based questions. They consider the HOTS problems are like the others. In addition, they assume that numeration-based literacy should be avoided because the learners are expected to understand the stories in the problems. Therefore, most learners have not had high-level thinking skills. That research aimed to develop a HOTS assessment instrument based on numeration literacy at the junior high school level.

Based on these conditions, educators should make assessments by giving HOTS-oriented problems based on numeration literacy (the ability to process information and knowledge using mathematically related numbers and symbols). Critical reasoning, together with problem-solving skills, presentation skills, creativity, innovation skills, and cooperation skills, are the elements of advanced thinking skills (HOTS) [7], [8]. HOTS fall into the category of high-grade thinking skills. HOTS are also used as a guide to learning cognitive thinking ability and psychosomatic in the 21st century [9]. The development of HOTS is linked to a psychosomatic capacity that requires knowledge, so they can be developed and enhanced through numeration literacy [10].

Assessments may be done using an instrument for numeration-based literacy. The numeration-based literacy also includes HOTS based on the minimum competence assessment. In minimum competence assessment, HOTS has become an activity in

developing and strengthening things through high levels of thought in various concepts [11].

HOTS is a strategic thinking ability using information to solve problems, analyze arguments, negotiate problems, or make predictions [12]. A high level would happen if a person stored information in memory and obtained new information, and then connected, arranged, and developed that knowledge to complete a goal or obtain answers that could solve a perplexing situation [12]. HOTS incorporate crucial reasoning, creative reasoning, problem-solving, and decision-making. Based on the above few suggestions, HOTS manage to be described as a high-level thinking skill that involves much of the process, including evaluating, thinking, developing arguments, applying principles to various contexts, writing, and creating.

High-level thinking skills are the reasoning system of learners at higher intellectual levels, from separate concepts, cognitive methods, and study taxonomy, such as problem-solving methods, Bloom taxonomy, learning, teaching, and evaluating [13]. Bloom's taxonomy is a conceptual frame to identify the thinking skills of the lowest levels up to the highest. According to Bloom, there are six levels of cognitive skills, among others: (C1) memory/remembering, (C2) perception, (C3) applying, (C4) analyzing, (C5) evaluating, and (C6) creation/creating [14]. Based on these levels, levels C1, C2, and C3 are classified as low-level thinking, while C4, C5, and C6 are classified as high-level thinking.

HOTS is very important for learning mathematics because they must use a broad mindset to find problems and allow students to make decisions or solve them [14]. One of the reasons why most learners have no HOTS is because teachers or educators do not explain the material concepts. They lack material, teach too quickly, and do not associate learning with everyday life [15]. Thus, the position of the teacher is influential in changing the mindset of learners during learning so they have higher-order thinking skills.

The competence to describe knowledge seriously so that everybody manages access to science and technology to upgrade their point of life is called literacy [16]. Literacy is the competence to read and, further, the capability to analyze and recognize the ideas behind writing. However, numeration ability is a person's competence to use mathematical knowledge to solve problems, explain events, or make decisions in everyday life [17]. Based on the opinions above, numeration literacy is the ability to use various types of numbers and symbols related to mathematics, analyze the information displayed, and generate predictions to assist in decision-making regarding everyday problems.

HOTS aims to improve students' abilities, especially in terms of critical thinking. Numeration literacy is very important for mathematical learning, just as is HOTS, because numeration literacy is one of the most helpful abilities in understanding mathematics in various fields, such as economics, engineering, science, social, and others [18]. There are three major obstacles that teachers or educators have found in developing the numeration literacy of learners. The special effort to directly work through numeration appears to be a nontraditional method of teaching mathematics [19]. Therefore, instruments or tools must contain numeration literacy to develop students' higher-level thinking skills.

The results showed that implementing the problem-based learning model results in different HOTS numeration literacy [20]. In addition, HOTS assessments in numeration literacy help students understand contextually based problems. HOTS are different from numeration literacy skills. Other studies have found that implementing problem-based learning in solving numeration literacy questions involved using learning tools and realistic scientific approaches [21]. In line with the research above, research tests the HOTS assessment instrument based on mathematical literacy on SPLDV material [4]. Many studies on the development of numeracy literacy-based HOTS assessment instruments

have been carried out, including the development of HOTS assessment instruments on SPLDV material [4] and the development of HOTS-based SPLSV questions [14]. However, among these studies, no research has developed a HOTS assessment instrument based on numeracy literacy for eighth-grade students in learning mathematics.

There is a problem in one of the junior high schools in Medan. A lack of explanations of the concept of HOTS in daily life prevents learners from resolving issues focused on HOTS that deal with everyday problems. Thus, this research aims to develop a HOTS-based assessment instrument on numeration literacy in mathematics subjects for eighth-grade junior high school students. HOTS-based numeration literacy assessment instruments also apply problems to daily life to make learners quick and alert to problem-solving and decision-making. Those factors coincide with the educational context of the industrial revolution 4.0. The study's novelty is the numeration literacy on mathematics materials that are more varied, unlike earlier studies focused on SPLDV material. This research also produced a collection of HOTS problems based on numeration literacy.

This research aimed to develop a HOTS assessment instrument based on numeracy literacy. The development of the instrument was carried out only on eighth-grade mathematics topics. On the other hand, previous studies have developed similar instruments but at different levels and with different subject matter. This study seeks to present a HOTS assessment instrument that stimulates higher-order thinking skills.

2. METHODS

This research uses the Research and Development method to create a HOTS-based numeration literacy assessment instrument to help students improve their higher-order thinking skills. Teachers and students can use the assessment instrument created as a guide in learning mathematics to achieve learning objectives.

The subjects of this research were limited to eighth-grade students at a junior high school in Medan, North Sumatra Province. The following steps were used to select research subjects: (1) purposively selecting an area, (2) purposively selecting a school in the area as a research location, and (3) selecting one of several existing classes to be the research subject. The selection of schools and classes made it easier for researchers because the demographic sample had been adjusted to the need for research. There were 30 students and a mathematics teacher who were interviewed.

The HOTS assessment instrument based on numeration literacy was developed using the 4D development model that consisted of 4 stages, namely defining, designing, and developing together, with dissemination [22]. They were used because they are systematic and organized. The structured stages of this model make it easier for the researchers to perform a procedure for assessing HOTS instruments to produce a product. The stages of the development can be seen in the figure below.

Based on the results of expert and practitioner validity tests, practicality tests, and field trials, revisions were needed until they were deemed fit for dissemination. The results of trials were used as a reference for evaluation and revision of the assessment instrument so that the final product could be produced from the development.

Data selection techniques contained (1) the measurement of the instrument, which contained HOTS questions and an assessment rubric; (2) validation of the contents of the instrument by teachers who teach mathematics and lecturers as expert validators to validate the designed instruments and student response questionnaires; (3) instrument revision on the validated draft by the validators; (4) field trials; (5) HOTS evaluation; (6) instrument revision based on the analysis of the trials.

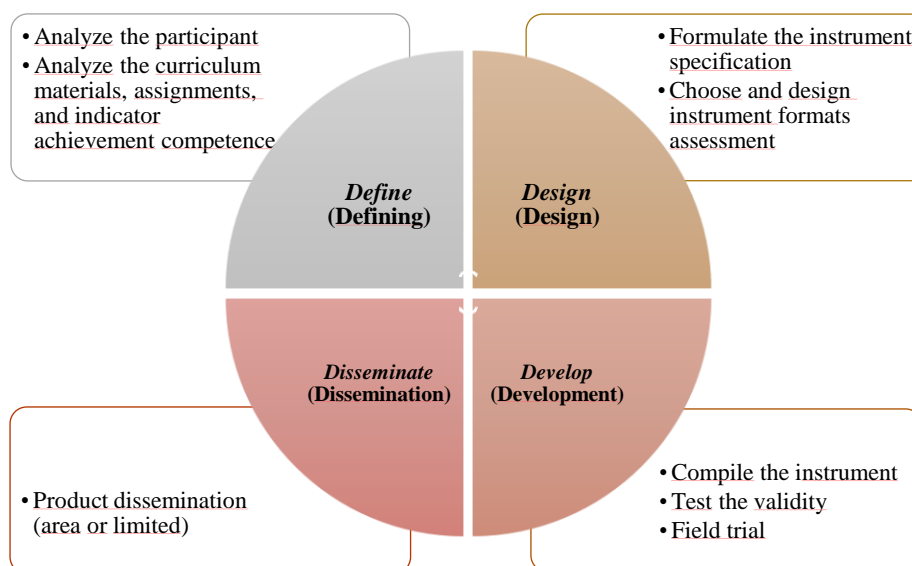


Figure 1. Research Procedure

This research produced some qualitative and quantitative data. Qualitative data was obtained from the validators' input, responses, criticism, and suggestions for improvement of the products. On the other hand, the quantitative data were obtained from validation expert assessment scores, teacher questionnaire scores, and student test scores in the field trials. The tests were conducted to measure the efficiency of the assessment instruments. The instruments aimed at measuring the validity, reliability, discriminating index, level of difficulty of the items, and practicality of the assessment instruments. The validity test included validation sheets by material experts together with media experts.

Product quality is very important and needs to be considered in research and development. To measure the quality of a product, it is reviewed based on five aspects: validity, reliability, discriminating index, level of difficulty of the items, and practicality. The data obtained was analyzed based on the answers to questions regarding the HOTS-based numeration literacy assessment instrument and whether it was valid, reliable, and practical. Based on this analysis, revisions were made.

Testing the instrument's validity was intended to ensure that the instrument was feasible. The validity test was carried out for each item using the Aiken V index. The data analysis was adjusted using the validity criteria listed in Table 1.

Table 1. Aiken V Index Validity Criteria [23]

SV in Percent (%)	Criteria	Information
$75\% < V \leq 100\%$	Valid	No revision needed
$50\% < V \leq 75\%$	Not Valid yet	Minor revision
$25\% < V \leq 50\%$	Not Valid yet	Major revision
$0\% < V \leq 25\%$	Invalid	Not applicable

After testing the validity of the items, the next step was to test the assessment instrument on students. Then, from the trials, the student's answers were obtained based on the assessment rubric to determine the validity, reliability, discriminating index, level of difficulty, and practicality. Validity is a measure that shows how valid an instrument is [22]. Based on the calculations with a significance level of 5%, if the values obtained are $r_{tabel} < r_{count}$, then the items of each instrument are said to be valid, and vice versa.

Furthermore, a reliability test was carried out to determine the accuracy of measuring the assessment instrument, meaning that the reliability test must produce the same measurement results even though it is carried out by different people, times, and places [24]. The reliability analysis employed the Cronbach Alpha formula.

$$r = \left(\frac{n}{n-1} \right) \left(\frac{s^2 - \sum pq}{s^2} \right) \tag{1}$$

Information:

- r : Instrument reliability
- n : The number of questions
- q : The proportion of subjects who answered incorrectly
- p : The proportion of subjects who answered correctly
- s² : Variance of total scores [25]

The results of the calculations were analyzed and adjusted according to the reliability criteria below.

Table 2. Reliability Criteria [25]

Coefficient Intervals	Criteria
0,00 – 0,50	Low
0,51 – 0,70	Moderate
0,71 – 0,90	High
0,91 – 1,00	Excellent

After carrying out the reliability test, the next step was calculating the discriminating index value. The item's discriminating index measures the instrument's ability level to differentiate students with high abilities and those with low abilities. The formula used to determine the discriminating index is as follows.

$$DP = \frac{BA - BB}{\frac{1}{2}N} \text{ or } DP = \frac{2(BA - BB)}{N} \tag{2}$$

Information:

- DP : Discriminating index
- BA : Proportion of participants who answered correctly
- BB : Proportion of participants who answered incorrectly
- N : The number of students who took the test

The calculation results can be analyzed based on the criteria listed in Table 3.

Table 3. Discriminating Index Criteria [24]

Difference Power	Criteria
0,71 – 1,00	Strong
0,31 – 0,70	Moderate
0,00 – 0,30	Weak

The difficulty level is identifying whether the questions are too easy or difficult. If the questions are too easy, they will not stimulate students' thinking. In contrast, If the questions are too difficult, they will cause students to lack enthusiasm. The formula used to determine the level of difficulty is as follows:

$$TK = \frac{\sum S}{N} \tag{3}$$

Information:

TK : Difficulty Level

$\sum s$: The number of students who answered correctly

N : Number of students

The results of calculations from the difficulty grade of each instrument item can be analyzed according to the subsequent table criteria:

Table 4. Difficulty Level Criteria [24]

Difference Power	Criteria
0,71 – 1,00	Easy
0,31 – 0,70	Moderate
0,00 – 0,30	Hard

After testing the difficulty level, the next step was to test the practicality based on the questionnaires' responses. This practicality test was used as a benchmark for a HOTS-based numeration literacy assessment instrument. The researcher gave a scale of 1 to 4 to the implementation of the assessment instrument. The following formula obtained the percentage of data:

$$Practicality \% = \frac{\text{Number of score each question}}{\text{Number of respondents}} \times 100\% \quad (4)$$

The practicality criteria are displayed in Table 5.

Table 5. Practicality Criteria [24]

Percentage (%)	Category
0,00 – 25,00	Impractical
25,01 – 50,00	Moderately Practical
50,01 – 75,00	Practical
75,01 – 100,00	Very practical

3. RESULTS AND DISCUSSION

The developed HOTS-based assessment instrument can improve students' higher-order thinking skills on circle material, circle tangents, and two-dimensional figures. This assessment instrument was developed using the 4D development model. This instrument was developed with the general aim of knowing the higher-order thinking skills of eighth-grade junior high school students in Medan. The specific objectives are to produce valid and reliable assessment instruments with good difficulty, differentiability, and practicality. The product was developed through a series of development processes, starting from the stages of definition (define), product design (design), product development (develop), and product dissemination (disseminate).

3.1 Definition Stage (Define)

This initial analysis was carried out by observing and obtaining problems. This stage began with observations and interviews with the mathematics teachers at one of the junior high schools in Medan. The researchers obtained information regarding the curriculum 2013 materials in school and the Merdeka curriculum. Based on the observations, it was revealed that students rarely use HOTS questions and rarely use questions based on numeration literacy. The instrument used by the teacher as an assessment of student learning outcomes was a test assessment.

Based on interviews with mathematics teachers, it was found that the instruments used were only essay questions without the HOTS question category. Therefore, the basic competencies and core competencies related to ongoing learning material are reviewed as an analysis of students' tasks. The researchers then examined the learning objectives by analyzing indicators of competency achievement. The analysis results from the interviews were also proven by looking at the students' scores on the daily mathematics test, especially the tangent to the circle. The data showed that 45% of students completed and 55% did not complete.

The questions were usually taken from the student handbook. The problem in the handbook was in the HOTS category, but they were not accustomed to answering using the HOTS procedures. The teacher used an assignment method that simply requires the student to understand what is being learned without setting out the completion steps of each problem. This finding indicated that students could not provide alternative solutions to problems, and their higher-order thinking skills were low. These problems motivated the researchers to develop an assessment instrument that can be used as evaluation material to improve students' higher-order thinking skills.

3.2 Design Stage (Design)

The design stage (design) aimed to produce an initial model (prototype) of the assessment instrument according to the learning material. This stage produced 50 multiple-choice questions based on HOTS and numeration literacy.

Table 6. Numeration Literacy Indicators

No	Numeration Literacy Indicator	Question Indicator
1	Students manage to use the ability to fix contextual problems in a common context.	Students are able to interpret problems in everyday life.
2	Students manage to interpret problems and solve them with formulas.	Students can interpret events or phenomena in the problem. Students can analyze problems using formulas.
3	Students use procedures useful in solving problems and can select strategies for solving problems.	Students can associate mathematics with everyday life. Students can assemble several geometric shapes in everyday life.
4	Students work efficiently with models and can select and combine different representations and then relate a problem to everyday life.	Students can interpret the meaning of the problem and choose how to solve the problem. Students can interpret the meaning of the problem and choose how to solve the problem.
5	Students manage tasks with models for complex situations and can solve complicated problems.	Students are able to project the data provided and prove unknown values. Students can project the data provided and prove how many things are needed to build a figure.
6	Students use thinking to solve numerical problems, make generalizations, and formulate and prepare their findings.	Students can interpret, design, and rearrange using formulas. Students can solve problems by compiling the parts requested.

After designing a question grid based on HOTS indicators and numeration literacy, the next step was adjusting the question grid with basic and core competencies. If the basic and core competencies align with the assessment instrument grid, then the researchers design the grid questions.

This stage was carried out by selecting and designing a format based on the results of student analysis and the curriculum. In the analysis, information was obtained that the student's mathematical abilities varied, and most of them were not familiar with

numeration literacy questions because they rarely got questions based on numeration literacy.

The researchers designed 50 multiple-choice questions. The questions were arranged in such a way according to the material chapter. The resulting questions consisted of 20 questions on circle material, ten on circle tangent material, and the remaining 20 on two-dimensional figures. Some questions in this assessment instrument were modified and adopted from questions found in mathematics books, provided that they met the PISA criteria and were included in the HOTS question category.

After compiling the questions, the researcher designed a validation sheet to validate them so that they were suitable for use as an assessment instrument and field trials. The researchers re-designed the teacher and student response sheets and the rubric of the assessment instrument.

3.3 Development Stage (Develop)

After designing the assessment instrument, the instrument was developed based on the formulated grid. The preparation resulted in a prototype called Design 1 which consisted of instrument identity, instructions, and 50 HOTS multiple-choice questions based on numeration literacy. Design 1 was then validated by the validator, followed by revisions to obtain improvements.

The assessment instrument went through two stages of assessment. First, the stage of testing the validity of the assessment instrument by two mathematics lecturers as validators. The results of the validator's analysis are as follows.

Table 7. Design 1 Validation Result

No	Aiken V index	Criteria
1 - 50	0.833 - 0.870	Valid

Based on the validation result by the validator using the Aiken V index, it was obtained that the HOTS-based numeration literacy assessment instrument consisting of 50 multiple-choice questions was declared valid and feasible to be used as a field trial assessment instrument. However, from the 50 approved questions, it was necessary to improve in terms of language structure and the writing of mathematical formulas.

The revised items were three questions on circle material, one on circle tangent material, and three on two-dimensional figures material. The mistakes of these questions lay in the lack of letters in the questions, fonts for formulas and mathematical units, and confusing sentences. Design 1, revised according to the validator's directions, resulted in Design 2. After design two was approved, the authors took ten questions randomly for field trials. The selected questions were converted into *essay format* to determine the students' HOTS-based numeration literacy.

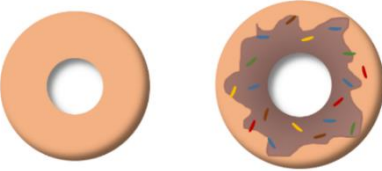
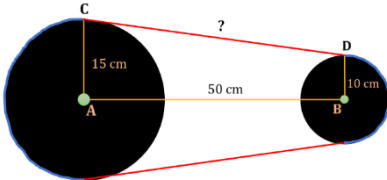
The problems originally developed in multiple-choice format were modified into description questions. The modification purpose was intended so that learners could write down the entire process of finding the correct answer but with varying steps. The numeration literacy competency has many ways to solve the problem. The constructed 50 multiple-choice questions have practical value. Therefore, the questions were included in the question bank as practice materials. The 10-question assessment instrument was re-validated by three subject matter and numeration literacy expert lecturers. The results of the second validation are as follows.

Table 8. The Result of the Second Validation

No Question	Aiken V index	Criteria
1 - 10	0.90 - 0.93	Valid

Based on the validation results using Aiken V Index, it was found that the HOTS-based numeration literacy assessment instrument consisting of ten description questions was declared valid and feasible for field trials. The following are examples of assessment instrument questions.

Table 9. Design 2 Assessment Instrument Development

No	Design 2
1	<p>Donuts are made from wheat flour, bread, yeast, sugar, egg yolks, and butter. Donuts generally have a characteristic, which is a circle and a hole in the middle like a ring. Unlike most bread types cooked by baking or steaming, donuts are fried.</p>  <p>Donuts are also one of Zahra's favorite foods. One day, Mother made donuts, and Zahra was happy to help her mother. Zahra, who was making donuts, printed the dough with an outer diameter of 7 cm and an inner diameter of 2 cm. However, the donuts expanded after frying by 0,5 cm. How wide is Zahra's donut after frying?</p>
2	<p>A motorcycle engine mechanic repaired a customer's motorbike that had a problem with a detached gear chain.</p>  <p>The mechanic checked the supply of spare parts in his workshop. It turned out that the chain with a 15 cm and 10 cm radius was out of stock. Suppose a pair of gears is placed at a distance of 50 cm from the center. How many meters of chain length does the mechanic need?</p>
3	<p>Engklek or hopscotch is a traditional game that uses simple tools to make. At least two people usually play Engklek. Physical education lessons are currently underway at school. Because there were still about 20 minutes left, Nurul and her friends decided to play hopscotch. After playing for a long time, Nurul and her friends felt tired and sat on a cube-shaped bench. There were five benches there. The benches were put together to form a cube. If each cube has a volume of $27.000cm^3$, what was the surface area of the benches?</p>

The field trials of the assessment instrument involved 30 students. The trial was carried out at one of the junior high schools in Medan. Data from the field trials was collected using questionnaires.

Table 10. Results of the Validity

No	Validity Index	No Question	Amount	Percentage
1	$\geq 0,349$ (Valid)	1,2,3,4,5,7,8,9,10	9	90%
2	$< 0,349$ (Invalid)	6	1	10%

Based on the analysis, there were nine questions with a percentage of 90% declared valid, and 1 question declared invalid. However, based on the results of the validation analysis by the validator, it was stated that all the items in the assessment instrument were

suitable for use. Other items declared valid can be used as an assessment instrument to train students' higher-order thinking skills and numeration literacy skills.

The reliability of the numeration literacy-based HOTS assessment instrument for ten essay questions was declared reliable with strong criteria, namely 0.983. Based on this finding, the HOTS-based assessment instrument can be developed as an assessment instrument. The following are the results of the discriminating index for each item.

Table 11. The Results of the Discriminating Index Analysis

No	Discriminating Index	No	Percentage
1	0,41 (Good)	8	10%
2	0,20 (Medium)	7	10%
3	0,01 – 0,13 (Weak)	1,2,3,4,5,6,9,10	80%

The outcome of the discriminating index analysis states that there was one item (10%) had a good discriminating index, one item (10%) had a moderate discriminating index, and eight items (80%) had a weak discriminating index. This test is a way to identify the groups of respondents by the index on the Rasch modeling [26] using analysis at the level of the individual's ability as a measure of whether or not learners can answer the question.

The quality of the overall instrument is good if the value of discrimination is higher because it can identify the classification of respondents and the question item [27]. Based on the outcome of the analysis, the discriminating index was measured to identify and distinguish students from the upper and lower groups. The analysis results of each item's difficulty level can be seen in the following table.

Table 12. The Results of the Difficulty Level Analysis

No	Difficulty Level	No Question	Total	Percentage
1	0,74 – 0,81 (Easy)	2,4,8	3	10%
2	0,33 – 0,66 (Medium)	1,3,5,6,7,9,10	7	80%

The analysis showed that three questions (30%) were at the easy difficulty level, and seven (70%) were at the moderate level. Based on the results of this analysis, the assessment instrument developed was at a moderate level of difficulty. The average level of difficulty for each item was 0.60 (medium). Based on this data, the difficulty level of the assessment instrument was in a good category. Based on the analysis, a good problem is relatively easy or difficult. Relatively easy problems are not encouraging learners to increase their problem-solving skills. In contrast, relatively difficult problems will cause learners to become discouraged at the task; thus lowering their spirits to do the task because they think the tasks are beyond their ability [28].

The practicality data of the assessment instrument was obtained using the Likert scale based on the teacher's and students' responses. The following is a table of the practicality of the assessment instrument.

Table 13. The Results of the Practicality Analysis

No	Respondents	Practicality Index	Criteria
1	Teacher	90%	Very practical
2	Students	67.25%	Practical

The assessment instrument used the HOTS aspects because it aims to unlock higher-order thinking skills [29]. Hence, the developed instrument of assessment must have standard-issue criteria. The criteria of the developed assessments are tested using product quality, validity, and practicality criteria [30].

The results of the analysis of the practicality of the questions on the instrument found that the instrument was very practical, with a percentage of 90% based on the teacher's responses and 67.25% based on the student's responses. Based on the results of this analysis, the assessment instrument was declared practical and suitable for use as an instrument for students to improve their HOTS and numeration literacy.

3.4 Stage of Dissemination (Disseminate)

This stage was carried out by disseminating the HOTS assessment instrument that has met the requirements. At this stage, the researchers disseminated the products in a limited way by distributing them as question banks in both softcopy and hardcopy to mathematics teachers at one of the junior high schools in Medan.

Based on the validity analysis using the Aiken V index, the 50 items obtained a value of 0.85. The same results were found in the ten description questions with an average value of 0.918. Therefore, both forms were declared feasible. Furthermore, the reliability test was done by comparing the R-table value's obtained value. If the variable value is higher than the R-table value, then the instrument has strong reliability, and vice versa. The analysis results above state that the questions are feasible to be used as a test instrument and are feasible to be tested directly on students. This study's results were compared with previous research results, which yielded feasible results [4].

Furthermore, a practicality test was carried out on the developed instrument using teacher and student response sheets. Three factors affected the practicality of the instrument: ease of implementation, ease of inspection, and clear instructions [31]. The results of the practicality test were obtained using teacher and student response questionnaires. A score of 90% was obtained from the teacher's responses which belonged to the very practical category, and a score of 67.25% was obtained from students' response questionnaires which belonged to the practical category. This study's results were compared with previous research results showing that the practicality test obtained an average score of 93.1 with very practical criteria [24]. Thus, it can be concluded that the practicality test on the assessment instrument was feasible to be used and practical. It is effective as a measuring tool for student learning outcomes.

Based on the results of field trials, it was found that several items did not meet the requirements. Therefore, design three was produced. Design 3 was a result of Design 2 revisions. Therefore, design 3 contained nine valid items to be part of the HOTS-based numeration literacy assessment instrument.

The field trials obtained a percentage of 87% from the HOTS question experts and 93% from the numeration literacy experts. The results of this development research are also supported by previous research, which developed a HOTS assessment instrument based on numeration literacy on SPLDV material with the highest validation score of 96% and categorized as very feasible [4]. Another research developed a mathematical literacy test instrument for junior high school students [32]. Numeration literacy differs from mathematical literacy. Mathematical literacy is not the same as numerical competence. The difference lies in the empowerment of knowledge and skill. This study includes the application of mathematical concepts and rules in real situations in daily life, coming up with many solutions, and dealing with non-mathematical factors. Mathematical knowledge alone does not preclude numeration.

Other research developed a HOTS assessment instrument but does not include numeration literacy [33]. This research is different from previous studies that contain only HOTS assessment instruments. According to some earlier studies, these studies link HOTS with numeration literacy in daily life and can be seen applying numeration literacy to the

HOTS instrument. Based on this fact, the development was carried out to meet the standard feasibility and practicality criteria for the learning process. The developed product is expected to develop and improve students' higher-order thinking skills in learning mathematics.

This research has limited materials. The highlighted material does not cover the entire eighth-grade materials. The material coverage is limited to a circle and two-dimensional figures. Furthermore, there are invalid problems because they do not meet the criteria. Hence, further research is expected to expand and revitalize the materials and questions intended to serve as instruments.

4. CONCLUSION

Based on the analysis, the developed HOTS assessment instrument based on numeration literacy in mathematics for the eighth-grade students of SMP/MTs produces nine feasible and practical essay questions. The findings can be proven from the HOTS problem and material experts' validation with a significance level of 5% with a percentage of 87% in the valid category. The percentage of 93% in the valid category was obtained from the validation by the numeration literacy experts. The instrument was declared reliable with very strong criteria (0.983). The discriminating index value of the questions was 0.11 in the weak category, and the average difficulty level value was 0.60 in the moderate category.

Furthermore, the practicality test through teacher and student response sheets obtained an average percentage of 67.25% in the practical category, and the practicality test by teachers obtained an average of 90% in the very practical category. Thus, the HOTS assessment instrument can improve students' higher-order thinking skills. In conclusion, the average students' scores obtained through the tests are categorized as high.

REFERENCES

- [1] Y. M. Cholily and et.al, "Pembelajaran di Era Revolusi Industri 4.0," *J. Ilm. Fak. Kegur. dan Ilmu Pendidik.*, vol. 6, no. 1, pp. 1–6, 2020.
- [2] M. Sukmoningsih, *Pengaruh Penggunaan Model Pembelajaran Realistic Mathematic Education (RME) terhadap Hasil Belajar Matematika pada Siswa Kelas IV Semester II SD N Kebowan 02 Kec.Suruh Kab. Semarang Tahun Pelajaran 2015-2016*. Salatiga : UKSW, 2016.
- [3] S. Maarif, *Analisis Kesulitan Belajar Matematika Siswa Kelas VII Madrasah Tsanawiyah Bahrul Huda Malang Materi Bangun Datar Tahun Ajaran 2019/2020*. Tulungagung : IAIN tulungagung, 2020.
- [4] N. Setyaningsih and T. D. Mukodimah, "Pengembangan Instrumen Asesmen Higher Order Thinking Skills (Hots) Berbasis Literasi Matematika Pada Materi SPLDV," *AKSIOMA J. Progr. Stud. Pendidik. Mat.*, vol. 11, no. 3, pp. 1739-1748, 2022.
- [5] U. Rosidin, N. Haryanti, H. A. Lora, and Viyanti, "Reconstruct the Class Assessment Strategy: Promoting the 21st-Century Learning," *Indones. J. Sci. Math. Educ.*, vol. 3, no. 1, pp. 27–36, 2020.
- [6] Y. R. Jonan, "Pengembangan Rubrik Penskoran pada Asesmen Otentik untuk Materi Volume dan Luas Balok," *J. Medives J. Math. Educ. IKIP Veteran Semarang*, vol. 4, no. 2, pp. 275-284, 2020.
- [7] Y. F. Kholifahtus, A. Agustiningsih, and A. A. Wardoyo, "Pengembangan Lembar Kerja Peserta Didik Elektronik (E-LKPD) Berbasis Higher Order Thinking Skill (HOTS)," *EduStream J. Pendidik. Dasar*, vol. 5, no. 2, pp. 143–151, 2022.

- [8] L. A. Purwasi and N. Fitriyana, "Pengembangan Lembar Kerja Peserta Didik (LKPD) Berbasis Higher Order Thinking Skill (HOTS)," *Aksioma J. Progr. Stud. Pendidik. Mat.*, vol. 9, no. 1, pp. 894–908, 2020.
- [9] T. S. Yen and S. H. Halili, "Effective Teaching of Higher-Order Thinking (HOT) in Education. The Online Journal of Distance Education and E-Learning," *TOJDEL Online J. Distance Educ. e-Learning*, vol. 3, no. 2, pp. 41–47, 2015.
- [10] Z. Arifin, "Pengembangan Instrumen Pengukur Berpikir Kritis Matematika Siswa SMA Kelas X," *J. THEOREMS (Orig. Res. Math.)*, vol. 1, no. 1, pp. 58–74, 2016.
- [11] Y. A. Pratiwi, R. C. Murti, and A. S. Nugraheni, "Analysis of Students Ability in Solving HOTS-Based Basic Indonesian," *AL-ISHLAH J. Pendidik.*, vol. 15, no. 1, pp. 687–694, 2023.
- [12] R. A. Sani, *Pembelajaran Berbasis HOTS (Higher Order Thinking Skills)*. Tangerang: Tira Smart, 2019.
- [13] D. R. Agustina, R. P. Wibawa, "Peran Pendidikan Berbasis Higher Order Thinking Skills (HOTS) Pada Tingkat Sekolah Menengah Pertama di Era Society 5.0 Sebagai Penentu Kemajuan Bangsa Indonesia," *Equilib. J. Ilm. Ekon. dan Pembelajarannya*, vol. 7, no. 2, pp. 137-141, 2019.
- [14] H. Maratin and A. Shodikin, "Development of Single Variable Linear Equation System Question Instruments Based on Higher Order Thinking Skill (HOTS)," *Pedagog. J. Ilmu Pendidik.*, vol. 22, no. 1, pp. 34–38, 2022.
- [15] R. P. Dalman and J. Junaidi, "Penyebab Sulitnya Siswa Menjawab Soal HOTS dalam Pembelajaran Sosiologi di Kelas XI IPS SMAN 1 Batang Kapas Pesisir Selatan," *Naradidik J. Educ. Pedagog.*, vol. 1, no. 1, pp. 103–112, 2022.
- [16] S. Dewayani and et.al, *Panduan Penguatan Literasi dan Numerasi di Sekolah*. Jakarta : Direktorat Jenderal Pendidikan Anak Usia Dini, Pendidikan Dasar, dan Pendidikan Menengah Kementerian Pendidikan dan Kebudayaan, 2021.
- [17] T. W. Liswati, Y. sri Yuniarti, and N. G. A. P. Sakinah, *Pengembangan Instrumen Penilaian Berbasis Literasi Numerasi*. Jakarta: Direktorat Sekolah Menengah Atas, 2019.
- [18] W. Han *et al.*, *Materi Pendukung Literasi Numerasi*. Jakarta Timur: Kementerian Pendidikan dan Kebudayaan, 2017.
- [19] A. Prabowo, S. Suparman, C. S. Li, D. Janan, and T. D. Damayanti, "The Effect of Reading Literacy to Mathematics Comprehension of Elementary School Students in Indonesia and Malaysia," *Int. J. Eval. Res. Educ.*, vol. 12, no. 1, pp. 546–554, 2023.
- [20] Y. Fitriyono, Rochmad, and Wardono, "Model PBL dengan Pendekatan PMRI Berpenilaian Serupa PISA untuk Meningkatkan Kemampuan Literasi Matematika Siswa," *Prism. Pros. Semin. Nas. Mat.*, vol. 4, no. 1, pp. 514–526, 2015.
- [21] A. Istiandaru, Wardono, and Mulyono, "PBL Pendekatan Realistik Saintifik dan Asesmen PISA untuk Meningkatkan Kemampuan Literasi Matematika," *Unnes J. Math. Educ. Res.*, vol. 3, no. 2, pp. 64–71, 2014.
- [22] Sugiyono, *Metode Penelitian Pariwisata (Kuantitatif, Kualitatif, Kombinasi, R & D)*. Bandung: CV Alfabeta, 2020.
- [23] D. Kristianti and S. Julia, "Pengembangan Perangkat Pembelajaran Matematika Model 4D Untuk Kelas Inklusi Sebagai Upaya Meningkatkan Minat Belajar Siswa," *J. MAJU*, vol. 4, no. 1, pp. 38–50, 2017.
- [24] N. M. Hadiawati, *Pengembangan Instrumen Penilaian Berbasis Literasi Numerasi menggunakan Quizizz untuk Mengukur Higher Order Thinking Skills (HOTS) pada Pembelajaran IPA di SMP*. Jember : Universitas Jember, 2023.

- [25] Waminton, *Evaluasi Hasil Belajar Matematika*. Yogyakarta: Media Akademi, 2015.
- [26] M. Erfan, M. A. Maulyda, V. R. Hidayati, F. P. Astria, and T. Ratu, "Analisis Kualitas Soal Kemampuan Membedakan Rangkaian Seri dan Paralel Melalui Teori Tes Klasik dan Model Rasch," *Indones. J. Educ. Res. Rev.*, vol. 3, no. 1, pp. 11-19, 2020.
- [27] B. Sumintono and W. Widhiarso, *Aplikasi Pemodelan RASCH Pada Assessment Pendidikan*. Cimahi: Trim Komunikata, 2015.
- [28] S. Arikunto, *Dasar-Dasar Evaluasi Pendidikan*. Jakarta : Bumi Aksara, 2015.
- [29] N. Izzati, R. Antika, N. Asma, and R. Siregar, "Pembimbingan Guru Dalam Mengembangkan Soal," *JMM J. Masy. Mandiri*, vol. 4, no. 3, pp. 370–381, 2020.
- [30] N. Nieveen and E. Folmer, "Formative Evaluation in Educational Design Research," *Design Research*, vol. 153, no. 1, pp. 152-169, 2013.
- [31] S. Arikunto, *Dasar–Dasar Evaluasi Pendidikan Evaluasi Pendidikan*. Jakarta: Bumi Aksara, 2018.
- [32] L. Auliya, *Pengembangan Instrumen Tes Literasi Matematis Berbasis Budaya Betawi untuk Siswa SMP/MTs*. Jakarta : Universitas Islam Negeri Syarif Hidayatullah, 2021.
- [33] R. Jaspita, S. Herawati, and Fauziah, "Pengembangan Instrumen Asesmen HOTS (Higher Order Thinking Skill) pada Pembelajaran Matematika Kelas VII Semester 1 SMP Negeri 23 Padang," *Seminar & Conference Proceedings of UMT*, vol. I, no. 1, pp. 1-6, 2020.