Development of the assessment model based on order thinking skills (HOTS) to measure students’ critical thinking

Nining Seyaningsih1 Widya Ayu Kurnia Sari1
1) Mathematics Education Department, Universitas Muhammadiyah Surakarta, Indonesia

Corresponding Email: ns259@ums.ac.id

Abstract. This study aims to develop the assessment model based on HOTS to measure students' critical thinking. This research uses the Research and Development (R&D) method. The instruments used in this study were assessment instruments and validation sheets of media and material experts. The characteristics of this study are the assessment instruments that measure the ability of students to solve problems that require high-level thinking skills and critical thinking. The subjects of this research are grade 8 students of SMP Muhammadiyah 4 Surakarta. The results showed that the assessment model developed was included in the good category used for the learning process. This is shown, the results of the validation of the material experts obtained percentage of the aspects of feasibility content is 0.92, percentage of the aspects of feasibility of construction is 0.82 and percentage of the aspects of language worthiness is 0.82. The results of the validation of the media experts showed a mean score of 3.6. The characteristics of instruments regarding distinguishing features and difficulty levels indicate the percentage of 0.39 and 0.49. Based on the results of student work using the HOTS-based assessment model in SMP Muhammadiyah 4 Surakarta shows that the ability critical thinking is still in the good enough category.

1. Introduction

Based on the education curriculum, mathematics is one of the subjects taught at schools in Indonesia. This is due to that elementary mathematics education is one of the main subjects that can form a critical, creative, logical and systematic mindset. The Students 'ability to think is still weak in mathematics because the delivery of material generally still uses a teacher-centered approach[1][2]. Mathematics is the right subject to develop high-level thinking patterns and hone students' critical thinking power. This is because mathematics is a systematic and structured science. To support this, one of the goals of mathematics learning is the ability to solve problems that are compatible with daily life. To realize the objectives of mathematics, it is necessary to do the assessment process.

Assessment becomes very important in the world of education. The assessment will be an evaluation material in accordance with the procedures used to obtain information to determine the level of ability and skills of students. The evaluating of the process and results of the teaching and learning activities is the goal in evaluating student’s outcomes by educators. The need for students to think critically and analytically according to international standards has been refined in the 2013 curriculum content standards, while the development of assessment instruments that measure thinking skills has been given space in the assessment standards. The cognitive processes are divided into two, according
to the revised Bloom Taxonomy, namely Higher Order Thinking Skills (HOTS) and Lower Order Thinking Skills (LOTS) [3][4]. The ability to remember, understand and apply is the ability to think that is contained in the ability to think at a lower level while the ability to analyze and synthesize, evaluate, and create or creativity is contained in higher-order thinking skills. High-level thinking ability (HOTS) is expected to be improved by assessing student learning outcomes.

The low ability to think becomes a problem that often arises in the process of developing assessment. The low ability to think can be caused by the lack of practice activities that hone the ability to think at a high level and learning activities that often give questions with the ability to think at a low level [5]. One of the problems is the lack of problems based on problem solving. In learning mathematics the ability to solve problems is an important component because it has an individual and practical role in life[2][6]. Further, Slamet et al all said that the reality in the field shows that the problem solving ability of students has not been well trained. In the process of learning mathematics students only memorize the knowledge given by the teacher and are less able to use that knowledge if they encounter problems in real life[2]. So it does not provide opportunities for students to improve their thinking skills.

This is as explained by Schoenfeld that factors causing low high order thinking skills are divided into two namely, external factors and internal factors[3]. The external factor is that the teacher gives more routine questions, mathematics learning in schools generally emphasizes routine problems that are questions that are structured in order to be presented clearly and contain all necessary information [7]. He also concluded that the error and obstacles factors in working on high order thinking questions are as follows: 1) lack of understanding and accuracy of students in the process of problem solving, 2) students' initial mathematical ability is low because the prerequisite material learned cannot be applied by students resulting in a learning process that is not optimal. This has an impact on the low ability of students in critical thinking.

Critical thinking is a thinking skill that must be developed, trained, practiced, and continuously integrated in learning[3]. Learning integrated with HOTS assessment can train students' thinking processes using mathematical logic. the use of HOTS assessments in mathematics learning is proven to effectively improve students' critical thinking skills, unlike HOTS assessments can train and develop important aspects of critical thinking skills [8]. From the research of Sumarni, et.l, stated that the assessment of the instruments developed gave different information about students' critical thinking skills. The combination of the two assessment schemes offers the potential for a variety of evaluations of students' conceptual understanding and critical thinking skills [9].

Based on the results of observations made by researchers at SMP Muhammadiyah 4 Surakarta, researchers found problems such as above. The lack of practice exercises that hone the ability to think at a high level and learning activities that often give questions with the ability to think at a low level. So that causes students in SMP Muhammadiyah 4 Surakarta is not accustomed to HOTS-oriented questions and questions that require solving critical thinking problems. Besides that, the results of PISA in the field of mathematics showed that in 2009 Indonesia was ranked 61st out of 65 countries, in 2012 Indonesia was ranked 64th out of 65 countries with a score of 375, in 2015 Indonesia was ranked 63rd out of 70 countries with score of 386 and in 2018 Indonesia was ranked 63rd out of 70 countries with a score of 386[10][11]. The PISA results showed that students' mathematical abilities are still low, when compared to other countries. The low PISA results are also supported by the results of the assessment from TIMSS, where in 2011, Indonesia ranked of 38 from 49 countries and in 2015 Indonesia was ranked 44th out of 49 countries[12].

Based on the background described above, it is necessary to have research to develop high-level thinking assessment instruments in the realm of critical thinking in mathematics subject matter in number patterns in junior high school students. Departing from these problems, it is necessary to develop a HOTS-based assessment model to measure students' critical thinking skills. By doing this research, teachers are expected to be able to improve critical thinking skills through the development of assessment.

2. Research Method
This research is a development research oriented to devise a product. The product is the assessment model based on HOTS. This study also attempts to develop assessment model based on HOTS to measure students' critical thinking. It was conducted at the one of Muhammadiyah Junior High School
in Surakarta, Central Java, Indonesia. Subsequently, the model generated from the development process was validated by an expert in mathematics education and mathematic teachers. This type of research is Research & Development, by adopting research according to Sugiyono [13]. As for the steps, namely: (1) the initial phase which includes looking at potentials and problems, gathering information and literature studies, (2) the product design phase which includes validating the design and revising the design, (3) the limited trial phase which includes and making revisions products that have been tested (4) the field trial phase and (5) the mass production of products phase[13]. The assessment instrument in the form of HOTS test items can be evaluated validated by the validation of the assessment instrument conducted by mathematicians. Mathematics education experts validate the product at the initial product development stage. The mathematical material developed by the number pattern from the grade 8, because this material can train students' thinking skills. According to Anno, learning number patterns can explore students' thinking abilities[14]. Marion in her research stated that completing number patterns can help students develop reasoning skills[15].

The sources of data in this study are: 1) experts in the field of model development for the validity of the model, 2) experts in the field of development, researchers and mathematic teachers for the practical assessment of the model, and 3) students as the participant in the implementation of the model for the measurement of the ability critical thinking[1].

The linkage of indicators of critical thinking with higher order thinking skills can be seen in Table 1 as follows.

Table 1. The Relation of the HOTS Indicators and Critical Thinking

| Indicators of Critical Thinking | Problem indicator | HOTS indicator |
|--------------------------------|------------------|---------------|
| 1. Focus on questions          | Presented a problem, rules, determine the main problem, criteria used to evaluate the quality, truth of an argument or conclusion. | Analysis: Assess, compare, criticize, sort, differentiate, determine, sort |
| 2. Analyze arguments           | Presented description of a situation: (1) concludes the argument quickly, (2) gives reasons that support the argument presented, (3) gives reasons not support argument presented | |
| 3. Make conclusions deductively | A statement that is assumed is assumed: (1) one conclusion that is true and logical, (2) two or more conclusions that are true and logical | Creating: Assembling, designing, designing, making, formulating. |
| 4. Make conclusions inductively | Presented a statement, information can determine an appropriate conclusion and give a reason. | |
| 5. Assess the definition        | Presented a description of a situation can determine: (1) positive and negative solutions, (2) which solution is the most appropriate to solve the problem presented | Evaluation: Evaluating, evaluating, criticizing, selecting / selecting, connecting, giving opinions |
| 6. Take a decision in action    | Formulate alternative solutions | |
The analysis technique is used to determine the validity of the model by recapitulating the experts’ statements and determining the validity by matching the results with the established criteria[1]. Meanwhile the practicality of the model is measured by recapitulating the experts’ statements and determining the practicality by embedding the results with the specified criteria. The scoring criteria used in this study are presented in Table 2.

| Criteria                        | Scoring Interval      | Category         |
|---------------------------------|-----------------------|------------------|
| **Validation for the material** | 0.80 < V ≤ 1.00       | Very high        |
|                                 | 0.60 < V ≤ 0.80       | high             |
|                                 | 0.40 < V ≤ 0.60       | High enough      |
|                                 | 0.20 < V ≤ 0.40       | Not high enough  |
|                                 | 0.00 < V ≤ 0.20       | Poor             |
| **Validation for media**        | 4                     | Very good        |
|                                 | 3                     | good             |
|                                 | 2                     | Good enough      |
|                                 | 1                     | Poor             |
| **Different power of the questions** | 0.00 - 0.19          | Not good         |
|                                 | 0.20 - 0.39           | Good enough      |
|                                 | 0.40 - 0.70           | Good             |
|                                 | 0.71 - 1.0            | Very good        |
| **Difficultly level**          | 0.00 - 0.29           | Difficult problem|
|                                 | 0.30 - 0.70           | Medium problem   |
|                                 | 0.71 - 1.0            | Easy problem     |
| **Critical Thinking skills**    | 90% < score ≤ 100%    | Excellent        |
|                                 | 75% < score ≤ 89%     | Very good        |
|                                 | 55% < score ≤ 74%     | good             |
|                                 | 40% < score ≤ 54%     | Good enough      |
|                                 | 0% < score ≤ 39%      | Poor             |

Table 2. Scoring Criteria
3. Result and Discussion

Development of the HOTS assessment instrument to measure critical thinking skills on the material pattern of the balloon refers to the research and development method [13] which consists of 4 phases of development namely the initial stage to look at potentials and problems and gather information, the product design phase, the limited trial phase and the field trial phase.

3.1 The initial phase

At the phase of seeing the potentials and problems, observations and interviews were conducted to see the potentials and problems regarding the needs and curriculum of the Muhammadiyah 4 Surakarta Junior High School which were conducted to the mathematics teacher and vice principal. Researchers also made observations to the library to find out what learning sources used in the junior high school, especially for class 8.

The results of seeing the potential that needs to be developed at the junior high school are, in mathematics learning, the assessment instruments used are still limited and inadequate so that it can be a potential for researchers to develop existing assessment instruments. The results of seeing the problems found in SMP Muhammadiyah 4 Surakarta, namely, students' ability in mathematics is lacking when viewed from previous test scores, students' ability to think that is still low and difficulty in understanding material number patterns, and the lack of students practicing working on problems train the level of thinking ability.

The problem is motivated by the lack of understanding of the material number patterns, the assessment instruments used do not contain questions that practice high-level thinking, and limited learning resources so that students have difficulty finding references to solving problems.

At the stage of gathering information researchers conducted interviews with mathematics teachers at SMP Muhammadiyah 4 Surakarta about what teaching materials were used in learning. The results of the interview are supplemented by the observation of the researcher. The results of the interview that the teacher only uses the textbook guidelines in the teaching and learning process. The teacher also explained that the lack of students' insights into HOT questions and the absence of HOTS assessment instruments used at school due to time and energy constraints in making them. In addition, the number patterns are material that is quite difficult for students to understand[5].

3.2. The Product Design Phase

In the product design stage, the researcher first analyzes the basic competencies and indicators. Next, the researcher determines the distribution of the questions and then arranges the questions based on the analyzed indicators. The questions compiled comprise the allocation of time used to solve problems, the form of questions and the relationship between the cognitive level and the level of critical thinking. The grids that have been compiled then form the questions. Following are the stages of designing a product

a. Analysis of Core Competencies, Basic Competencies and Material Indicators

The analysis phase is carried out to assess the extent of the material to be used in the preparation of the instrument. In this case, researchers are guided by the Minister of Education and Culture Regulation No. 21 of 2016 concerning content standards and Number 24 concerning core competencies and basic competencies. Basic competencies and indicators of achievement of competencies in material number patterns can be seen in Table 3 below.
Table 3. Basic competencies and material indicators of number patterns

| Basic Competency | Competency Achievement Indicator |
|------------------|---------------------------------|
| 3.1              | Make generalizations of patterns in rows of numbers and configuration of object configurations |
|                  | 3.1.1 Identify integer patterns. |
|                  | 3.1.2 Apply patterns and generalizations to make predictions |
|                  | 3.1.3 Conduct experiments to generalize number patterns or object configurations. |
|                  | 3.1.4 Students can find links between syllable patterns or numbers in object configurations. |
| 4.1              | Solve problems related to patterns in sequence numbers and configuration configurations of objects |
|                  | 4.1.1 Solve problems related to number patterns, sequences, sequences |
|                  | 4.1.2 Uses to solve real problems |

Table 4. Instrument Assessment Feasibility

| Aspect Assessed   | No | Validator                  | Average Score | Category   |
|-------------------|----|----------------------------|---------------|------------|
| Content / Material| 1  | Material Expert (Lecturer) | 0.96          | Very good  |
|                   | 2  | Material Expert (Teacher)  | 1             | Very good  |
Based on Table 4, the three material expert validators stated that the appropriateness of assessment instruments in the category is very feasible, then it can be concluded that the appropriateness of the assessment instruments from the aspect of content, construction and language obtain the results of the very feasible category assessment.

Table 5. Results of the assessment of the appropriateness of the overall assessment instrument

| No | Validator                  | Total score | Average score | Category  |
|----|----------------------------|-------------|---------------|-----------|
| 1  | Material Expert (Lecturer) | 345         | 0.84          | Very high |
| 2  | Material Expert (Teacher)  | 407         | 0.98          | Very high |
| 3  | Material Expert (Teacher)  | 369         | 0.90          | Very high |
|    | total                      | 1121        | 0.91          | Very high |

Based on Table 5, the results of the validation by the material expert validator can be concluded that the HOTS assessment instrument to measure the critical thinking ability of junior high school students in class 8 in the first semester in the number pattern material was declared suitable for use in the learning process.

The results of the HOTS assessment instrument validation to measure the critical thinking skills of the media expert validator are presented in Table 5 as follows.

Table 6. The results of assessment instrument validation by media validators

| No | Assessment aspects | Total score | Average score | category  |
|----|--------------------|-------------|---------------|-----------|
| 1  | General display    | 21          | 3.5           | Very good |
| 2  | Special display    | 11          | 3.7           | Very good |
| 3  | Media presentation | 15          | 3.8           | Very good |
|    | Average            | 47          | 3.6           | Very good |

Based on Table 6, the results of media validation by the media validator expert assessment instruments show very good categories. Then the HOTS assessment instrument to measure the critical thinking skills of junior high school students in class 8 in the first semester in the material of number patterns is appropriate to be used / tested.

3.3. The Limited Trial Phase

The next step is the limited trials phase. This phase was conducted to determine the appropriateness of the assessment instrument when viewed from the level of difficulty and distinguishing features of the questions. The limited trial phase the researcher obtained the results of student work on the items on the assessment instrument given from three students. Analysis of the results of the limited trial shows the results for multiple choice have a level of difficulty that includes the medium category and
distinguishing power in either category. As for the description shows the level of difficulty in the medium category and distinguishing power in the very good category.

This trial is carried out by giving questions to students and done on the answer sheets provided. The instrument provided were 25 items, with 20 multiple choice questions and 5 essay questions. The limited trials were conducted to determine the appropriateness of the assessment instrument when viewed from the level of difficulty and distinguishing features of the questions. Analysis of the results of the limited trial shows the results for multiple choice question have a level of difficulty that includes the medium category and distinguishing power in either category. As for the essay questions show the level of difficulty in the medium category and distinguishing power in the very good category.

Difficulty and differentiation analyzes that have been carried out on multiple choice items and descriptions will be used to determine the quality criteria of the HOTS assessment instrument whether or not it is appropriate to be widely tested. Based on the results of the analysis of the level of difficulty and distinguishing power obtained conclusions, namely:

1) The assessment instrument has a difficulty level of 0.43 for the multiple choice questions and 0.55 for the essay questions so that it is declared in good category.
2) The assessment instrument has a degree of distinguishing power of 0.47 for the multiple choice questions and 0.31 for the essay questions, so that it is declared in good category.

So the HOTS assessment instrument to measure critical thinking skills can be tested widely because it has a good instrument criteria and can be tested.

Based on the limited trials conducted, the assessment instruments that have been tested are limited then the level of difficulty and distinguishing power are analyzed. The analysis results that have been obtained by researchers indicate that the HOTS assessment instrument has a good category and can be tested widely. So at this stage the researchers only made a few revisions to the HOTS assessment instrument on the use of conjunctions and spellings that still contained errors. However, the revision of phase did not change the instrument in content.

3.4. The Field Trial Phase

The field trial phase the researcher obtained the results of student work on the items on the assessment instrument given from 17 students. The results of the HOTS question work showed that there were 5 students in the category, only 6 students were in the good category and there were 6 students in the good category. Whereas in working on critical thinking questions, it shows the ability of students to think critically that is included in the category of very less there are 5 students, the category lacks 8 students and there are only 8 students enough categories. In brief, it can be seen in Table 7. below.

| No | Item Type Question | Critical Thinking | |
|----|--------------------|--------------------|---|
| 1  | Multiple choice    | 135                | 18 |
| 2  | Essay              | 251                | 15 |
| 3  | Score              | 51                 |    |
| 4  | Category           | Good enough        |    |

From Table 7, it can be seen that the eighth grade students of SMP Muhammadiyah 4 Surakarta have high-level thinking skills in the sufficient category and have critical thinking skills in the less category. The assessments model based on HOTS can make important contributions to improve students' critical thinking skills including problem-solving skills, decision making, conclusions, divergent thinking skills, evaluative thinking skills, and creative skills. What if students' critical thinking skills
can be achieved well. This is expected to improve the quality of learning, especially to master the mathematical concepts learned. Because HOTS questions can be solved in more than one way, students are therefore required to think critically in determining how to solve problems. So the application of this assessment instrument will improve the ability to think at a higher level and the ability to think critically.

This research is in line with Wayan Wardana’s research that HOTS assessment can make contribution to improving students' critical thinking skills including problem-solving skills, decision making, evaluative thinking skills, and creative skills [8]. The research about critical thinking was conducted by W Sumarni in the Development of assessment instruments to measure critical thinking skills, that the assessment instruments developed have met the criteria for good test quality, so that it can be used as an alternative assessment instrument for teachers to reveal students' critical thinking skills[9]. Two assessment schemes developed based on different approaches, therefore, can provide different information about students' critical thinking skills. Then the application of this assessment instrument will improve the ability to think at a higher level and the ability to think critically.

The design assessment model has been developed to develop from the theories that already exist. The assessment instruments that have been prepared have met the eligibility and effectiveness criteria for use in the learning process. So after the field trial phase, the assessment model based on HOTS to measure the critical thinking skills of students of class 8 of junior high school in the semester in the material of number patterns can be mass produced.

4. Conclusion
The research and development that has been carried out produces a product that is the HOTS assessment instrument to measure the critical thinking skills of students in class 8 of number pattern material. The resulting product design includes the results of the analysis of core competencies, basic competencies and material indicators, a grid of HOTS questions and critical thinking, and the development of HOTS questions and critical thinking.

The results of the validation from the material experts obtained the percentage of the feasibility of the content / material, the feasibility of the construction aspects and aspects of the feasibility of the language that is 0.92, 0.82 and 0.82. The results of the validation of the media experts showed a mean score of 3.6. In the product trial process, instrument characteristics regarding distinguishing features and difficulty levels showed a percentage of 0.39 and 0.49. Thus the assessment instruments that have been developed are included in the good category for use in the learning process. The results of the HOTS assessment instrument field test to measure critical thinking skills showed a value of 0.51 which was included in the good enough category. From the steps that have been carried out, it can be concluded that the assessment model based on HOTS is feasible to use and can measure the ability to think critically using HOTS questions.

References
[1] Setyaningsih N and Sri Rejeki N 2019 Developing Realistic and Child-friendly Learning Model for Teaching Mathematics J. Res. Adv. Math. Educ. Vol. 4, No 15–24
[2] Slamet N S 2010 Pengembangan Materi dan Model Matematika Berbasis Media dan Berkonteks Lokal Surakarta Dalam Menunjang KTSP
[3] Lewis, A. & S 1993 Defining Higher Order Thinking Skills (Theory into Practice, 32 (3))
[4] Abosalem Y 2016 Assesment Techniques and Student Higher Thinking Skills Int. J. Second. Educ. 4(1) 1–11
[5] Nining Setyaningsih 2013 Implementasi Model Problem Based Learning Dengan Pendekatan Scientific untuk Meningkatkan Pemahaman Konsep Matematika Siswa (Surakarta Indonesia)
[6] Barnes H 2004 A development case study : Implementing the Theory of Realistic Mathematics Education with low attaining leaners (University of Pretoria, South Africa)
[7] King, F.J. & R 2006 Higher Order Thinking Skills. Center for Advancement of Learning and Assessment
[8] Wardana I W 2018 Higher Order Thinking Skills Assessment towards Critical Thinking on Mathematics Lesson Int. J. Soc. Sci. Humanit. 2 1–13
[9] W Sumarni, K I Supardi N W 2018 Development of assessment instruments to measure critical thinking skills IOP Conf. Ser. Mater. Sci. Eng. Volume 12 hal 24-40
[10] OECD 2018 Result in Focus vol OECD Publi
[11] OECD 2016 Result in Focus vol OECD Publi
[12] Puspendik 2015 Hasil TIMSS 2015 Diagnosa Hasil untuk Perbaikan MuUt dan Peningkatan capaian
[13] Sugiyono 2009 Metode Penelitian Kuantitatif, Kualitatif dan R&D (Bandung: Alfabeta)
[14] M A 1983 Anno’s Mysterious Multiplying
[15] Marion 2015 Desain Pembelajaran Pola Bilangan Menggunakan Jaring Laba-laba di SMP J. Kependidikan 45 No. 1
[16] Heong, Y.M., Othman, W.D., Md Yunos, J., Kiong, T.T., Hassan, R., & Mohamad M. 2011 The Level of Marzano Higher Order Thinking Skills Among Technical Education Students Int. J. Soc. Humanit. 1(2) 121–125.