Al-Ishlah: Jurnal Pendidikan
Vol.14, 4 (December, 2022), pp. 5393-5400
ISSN: 2087-9490 EISSN: 2597-940X, DOI: 10.35445/alishlah.v14i4.1411

Development of Mathematics Assessment Instruments for Senior High School Based on Higher-Order Thinking Skills (HOTS)

Ellis Mardiana Panggabean[1], Feri Haryati[2], Sri Wahyuni[3]

[1] Universitas Muhammadiyah Sumatera Utara, Indonesia; eliismardiana@umsu.ac.id
[2] Universitas Muhammadiyah Sumatera Utara, Indonesia; ferihayati@umsu.ac.id
[3] Universitas Muhammadiyah Sumatera Utara, Indonesia; sriwahyuni@umsu.ac.id

ARTICLE INFO

Keywords:

Higher Order Thinking Skills
HOTS;
Mathematics Assessment;
Senior High School

Article history:

Received 2021-08-03
Revised 2022-01-10
Accepted 2022-11-27

ABSTRACT

This study aims to develop a mathematical assessment instrument based on higher order thinking skills. This research is adevelopment research that aims to develop a valid, practical and effective HOTS assessment instrument. The development model used is the Tessmer model. In conducting this research, the researchers adhered to the five steps applied in this study, namely preliminary, self-evaluation, expert review, one-to-one and small group and field testing. The findings of this study reported that these tested items were appropriate to be used as alternative teaching material for senior high school students who learned Mathematics in order to develop their higher-order thinking skills since each item meets the validity of content (content), construct and language. In conclusion, the development of these assessment instruments was considered successful.

Corresponding Author:
Ellis Mardiana Panggabean
Universitas Muhammadiyah Sumatera Utara, Indonesia; eliismardiana@umsu.ac.id

1. INTRODUCTION

Logical thinking skills, critical thinking skills and reasoning abilities are the scope of Higher Order Thinking Skills (HOTS). HOTS is thinking at a higher level than remembering facts or stating a fact. These basic skills are needed in everyday life and are part of academic achievement in schools (Marshall & Horton, 2011). To improve student performance, the nature and development of these skills, especially in learning mathematics, must be understood by all teachers in all education sectors (Santos-Trigo & Moreno-Armella, 2013). The importance of using HOTS in learning is also stated by Pratama & Retnawati (2017) that HOTS is a concern in education and is a partner for 21st-century skills. As is known, the skills needed in the 21st century are *critical and creative thinking*. Critical and creative thinking is one of the variables that can increase student achievement. The findings of both show a positive correlation between *higher-order thinking skills* and student learning outcomes.

Higher-order thinking skills are built with specific learning strategies in a learning condition, the paradigm of intelligence as a system, a shift in view towards multidimensional and interactive, and more specific thinking skills (Chun & Abdullah, 2019). If viewed from the point of view, the assessment based on higher-order thinking skills aims to improve students' thinking skills at a higher level. Especially concerning the ability to think critically in receiving various types of information, think creatively in solving a problem by using the knowledge possessed, argue well, and construct explanations and make decisions in complex situations. Through HOTS, students are expected to learn things they do not know and then successfully apply them to new situations. These abilities are indeed very much needed by the younger generation to face the Industry 5.0 era, which has uncertain work dynamics (Saputro et al., 2020). An environment with various types of problems and different human origins demands to be easy to adapt so that the ability of HOTS is very supportive and valuable.

The extent to which mathematics teachers understand the role of HOTS in learning mathematics is a question. Hinde & Perry (2007) commented that teachers generally link cognitive skills with the school curriculum. Attention is limited to the role of these skills in facilitating student thinking and learning in specific subjects. Furthermore, teachers have limited opportunities to engage in activities with students that will support the development of higher-order cognitive abilities. The ability to integrate HOTS into mathematics education can be hampered by factors, including teachers' limited knowledge of HOTS. Thus, the development of-based mathematics assessment instruments higher-order thinking skill (HOTS) needs to be done because the HOTS-based mathematics learning assessment instrument can help prepare students for the 21st century learning, increase students' learning motivation, and is a reliable source of information for making decisions about learning. a

Thomas & Thorne (in Nugroho, 2018) state that higher-order thinking is thinking higher than simply remembering facts or presenting facts or applying rules, formulas and procedures. Higher-order thinking requires someone to do something about facts, i.e. understand them, infer them, relate facts in new ways and apply them to solve problems. Higher-order thinking occurs when students acquire new knowledge and store it in memory, and then this knowledge is correlated with previous learning to achieve specific goals. In other words, HOTS occurs when someone gets new information, holds it, organizes and associates it with existing knowledge and then disseminates that information to reach particular objects or solutions to a problem. At higher-order thinking, students will tend to use logic rather than just remembering and memorizing formulas, so students will master concepts and solve more complex mathematical problems.

Higher-order thinking requires students to manipulate information and ideas by transforming their understanding and implications. This transformation will occur if students combine facts and ideas by synthesizing, generalizing, explaining, formulating hypotheses, concluding or interpreting. Manipulating information and ideas through this process allows students to solve problems and find new understandings for themselves. If these indicators are synergized with Bloom's taxonomy, the HOTS indicators that can be used are; 1) Analysis Level; 2) Evaluation Level; and 3) Creating level. The following is a picture of the building of *higher-order thinking skills* based on the revised edition of Bloom's taxonomy.

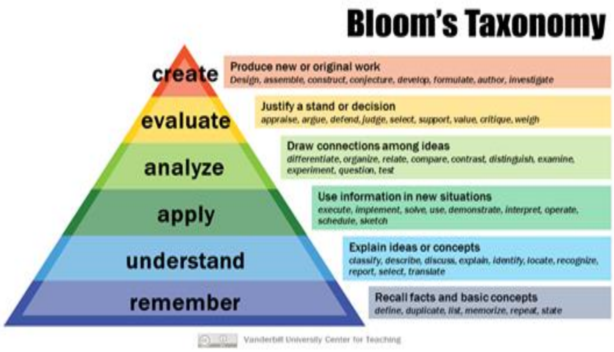

Fig 1. Image of building *Higher Order Thinking Skills*

Analyzing is breaking down (decomposing) the material into its constituent parts and then determining the relationship between legs and as a whole. The analysis level consists of the ability to distinguish, organize and relate. Evaluation level is the ability to make decisions based on established criteria. This level consists of checking and critiquing skills. Checking or rechecking is a process to find inconsistencies in processes or products. Criticizing is the process of assessing an opinion or result based on a set of predetermined criteria.

A study of Panggabean et al. (2020) shows that the test items compiled and used by teachers in evaluating cognitive mathematics learning outcomes mainly emphasize low-level thinking with easy answers. At the end of the report, it is suggested that the assessment instrument for student achievement be changed because students are required to have new skills and knowledge adapted to the era of the 21st-century global economy wherein this era students not only need basic understanding but also need to think critically, analyze and make appropriate decisions based on the results of the analysis. Therefore, this study is conducted. At creating, students organize various information using new or different ways or strategies than usual. Students are trained to combine the parts to form something new, coherent, and original. Based on the characteristics above, the HOTS-based assessment instrument developed and built-in this study uses HOTS indicators, namely (1) identifying and connecting relevant information from the problem; (2) making accurate conclusions based on the information obtained; (3) finding consistency/inconsistency in the product; (4) evaluate the product against the specified criteria/standards; (5) synthesize ideas/problem-solving strategies; (6) apply problem-solving strategies; and (7) developing new alternatives for problem-solving. The novelty in this study lies in the HOTS-based test, whose material is flat-plane analytical geometry.

2. METHODS

The development model used in this study is the Tessmer (1993) model. This study applies five steps: preliminary, self-evaluation, expert reviews, one-to-one (low resistance to revision) and small group and field tests (high resistance to correction).

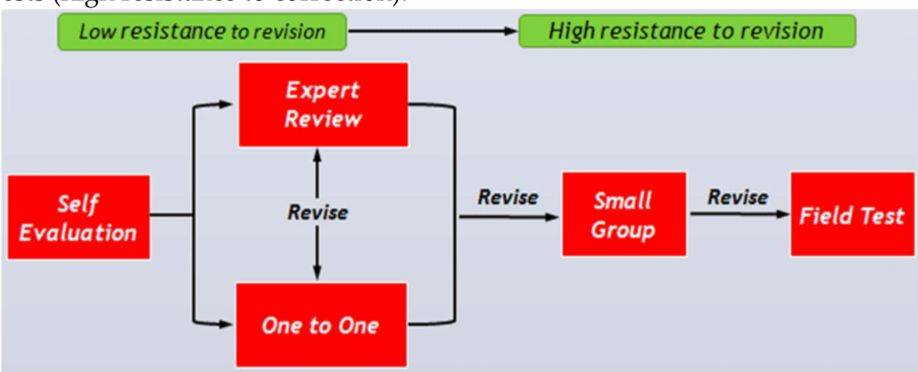

Fig 1. Tessmer's model of development

This preliminary stage consists of two parts, namely the preparation stage and the design stage. At this preparatory stage, curriculum analysis, research subject analysis and research scheduling were carried out. The design stage was carried out by compiling a table of specifications for the flat analytic geometry test material, including KD, material, question indicators, and cognitive level. Based on the table, the questions were designed. Designing questions as an assessment instrument was carried out by prototyping, namely the analysis, planning, and implementation phases simultaneously and repeatedly. The design results were given the name *prototype*, a model or basic description of the product, and initial testing. The assessment of each *prototype* focused on the characteristics of validity in the form of content (content), constructs and language. Content validity was assessed by comparing the suitability between the questions and the indicators set for each material. If the query matches the hand based on the expert's assessment, the question has helpful content. Otherwise, if the expert judges that the question deviated from the indicator, the question was declared content invalid. Construct validity refers to the extent to which an instrument measures the concept of a theory that forms the basis for the instrument's preparation. This study relates to the suitability of the questions designed with the theory and criteria for HOTS questions. Furthermore, the validity of the language was assessed for the language used in the questions with the correct Indonesian language rules and did not contain double meanings or ambiguities.

3. FINDINGS AND DISCUSSION

3.1 Research Findings

The preparation of *Higher Order Thinking Skills* (HOTS) questions or higher-order thinking skills requires various criteria, both in terms of the form of the questions and the content of the subject matter. The technique of writing good HOTS questions in the form of multiple-choice and descriptions is generally the same as writing low-level questions. Still, several characteristics distinguish them. The author used several ways as guidelines to write questions that require high-level thinking, namely, the material being asked is measured according to Bloom's cognitive domain. At the beginning of each question, a stimulus was given in sources/reading materials such as pictures, graphs, formulas, tables and examples. Before designing the questions, a curriculum analysis was carried out. Here are the results of the curriculum analysis that the author has carried out.

The course analysed was the Analytical Geometry of Flat Fields. The material in this course contains studying two-dimensional coordinate systems, vectors in planes, straight line, circle, parabola, ellipse and hyperbola. The Basic Competency is after studying this course, second-semester students of the Mathematics study program are expected to be able to understand, to draw graphs of points on a Two-Dimensional Coordinate System, Vectors in the Field, formulate equations of straight lines, circles, ellipses, parabolas, hyperbolas and their graphs, and be able to solve daily problems related to it.

Based on the analysis of these essential competencies, the authors formulated the indicators that, of course, are included in the high level, then composed a test specification table (attached), chose an exciting stimulus and then arranged the items according to the indicators.

Table 1. Outline of Lecture

No.	Subjects	Sub-topics
1	Cartesian CoordinateCartesian Coordinate	SystemSystem Distance Two points on a plane. Polar Coordinates of
2	Vectors in a	Vector Plane in a Long PlaneVector AngleFormed by Two Vectors Operations on a Vector
3	Straight Lines	Equation of a straight line The position of two lines on a plane An angle formed by two lines Distance of a point from a line Line Beam
4	Circle	Equation of a circle Tangent to a circle File circle
5	Parabola	Equation of parabola Equation of tangent to parabola
6	Ellipse	Equation of ellipse Equation of tangent to parabola

		Equation of hyperbola
7	Hyperbola	Equation of a tangent to hyperbola

Tests based on *higher-order thinking skills* designed in this research are in multiple-choice with five answer choices. The number of question items is 40, with C4 as 18 items, C5 as 13 items and C6 as nine items. The process of designing questions as an assessment instrument is carried out by *prototyping*, as stated in the previous chapter (method). The design results are named *prototype*, a model or basic description of the product for initial testing.

Assessment of each *prototype* focuses on the validity of the content, construct and language. Content validity is assessed by comparing the suitability between the questions and the indicators set for each material. If the query matches the indicator based on the expert's assessment, the question has helpful content. Otherwise, if the expert judges that the question deviates from the indicator, the question is declared content invalid. Construct validity refers to the extent to which an instrument measures the concept of test indicators *higher-order thinking skills.* Furthermore, the fact of the language is carried out in questions following the correct Indonesian language rules and does not contain double meanings or ambiguities.

Based on the results of the validity test by 5 (five) experts (*expert reviews*), it was concluded that: 1) Question Items 1 to 40 had met both contents and construct validity; 2) In terms of language and question writing, all questions are stated as very understandable (SDP); 3) several questions need revision, namely minor revisions and significant revisions. Those that need minor revisions are items 1, 3, 30, 33 and 37. The following is a table of items that need minor revisions and the results of the revisions.

Table 2. Question Items That Need Minor Revision

No.	Before Revision	After Revision
1	A university student stated that he was standing at a place that he ***claimed*** had the coordinates (2, k) measured from a place he considered the origin of O (0,0). His two other friends are standing elsewhere with coordinates (3,7) and (9,1). If the distance between the student and each of his friends is the same, then k is worth a. 7 b. 1 **c. 0** d. 2 e. 3	A university student states that he is standing somewhere ***with the coordinates*** (2,k) measured from a place he considers the origin point O(0, 0). His two other friends are standing elsewhere with coordinates (3,7) and (9,1). If the distance between the student and each of his friends is the same, then k is worth f. 7 g. 1 **h. 0** i. 2 j. 3

2		
	2. ***Given*** the polar coordinates of the two cities, P(r_1,θ_1) and Q(r_2,θ_2), the distance PQ = d is **a.** $d=\sqrt{r_1^2+r_2^2-2r_1r_2\cos(\theta_1-\theta_2)}$ b. $d=\sqrt{r_1^2+r_2^2-2r_1r_2\cos(\theta_1+\theta_2)}$ c. $d=\sqrt{r_1^2+r_2^2+2r_1r_2\cos(\theta_1-\theta_2)}$ d. $d=\sqrt{r_1^2+r_2^2+2r_1r_2\cos(\theta_1+\theta_2)}$ e. $d=\sqrt{r_1^2+r_2^2+2\cos(\theta_1-\theta_2)}$	***Given*** the polar coordinates of the two cities, P(r_1,θ_1) and Q(r_2,θ_2), the distance PQ = d is **a.** $d=\sqrt{r_1^2+r_2^2-2r_1r_2\cos(\theta_1-\theta_2)}$ b. $d=\sqrt{r_1^2+r_2^2-2r_1r_2\cos(\theta_1+\theta_2)}$ c. $d=\sqrt{r_1^2+r_2^2+2r_1r_2\cos(\theta_1-\theta_2)}$ d. $d=\sqrt{r_1^2+r_2^2+2r_1r_2\cos(\theta_1+\theta_2)}$ e. $d=\sqrt{r_1^2+r_2^2+2\cos(\theta_1-\theta_2)}$

In trials one-one and small groups with five students to determine the level of readability of each item found that all students understood the questions well enough. The trials conducted on students who have programmed the Field Analytical Geometry Course found unexpected results. For each basic competency, the average acquisition percentage is below 50%. These results indicate that students' higher-order thinking skills are still low. The descriptive analysis of student responses showed 100% happiness with the prepared teaching materials, the readability of teaching materials can be understood (90%).

3.2 Discussion of Research Results

The results of content validation by experts stated that all the questions compiled met the valid criteria. According to the researcher, this is caused by the substance of presenting the competencies being assessed. The planning step determines the appropriate competencies to be evaluated, compiling indicators of learning outcomes based on the competencies formulated and others fulfilled (Ramadani et al., 2017). If you pay attention again, the answer choices are homogeneous and logical in terms of material. The language used can be understood by students.

The characteristics of the HOTS-based items are distinguished by the form and content of the subject matter (Misri et al., 2020; Widiana, 2020). In general, writing multiple-choice questions is the same as writing essay questions and LOTS (low order thinking skills). The guideline used as a guide for the author was that the materials' contents in the items measured by behaviour according to Bloom's cognitive level taxonomy (Tarman & Kuran, 2015). Each item is given questions that measure higher-order thinking skills. The questions are arranged in cases, mathematical equations, formulas, pictures, graphs and tables.

In preparing these items, the provisions for making HOTS questions and examples of questions indicators were very well understood. Likely, they focused on the questions. For example, the problem indicator was that data presented (can be in the form of rules, pictures, graphs or tables), so students could test the truth of the argument or draw conclusions based on the data (Höst et al., 2000). Items that analysed statements, the indicator of the question was that a description of a situation or one/two arguments presented. Students concluded the arguments correctly from some of the opinions given. To compare a conclusion, the question indicator was provided with a statement assumed to be accurate, while the choice consisted of a correct and logical conclusion. In this case, students can choose findings that follow the ideas presented. Characteristics of the items assessed in the indicator question have given a description of a situation, a problem statement, and the possibility of solving the problem. Students can determine positive and negative solutions or which solution is the most appropriate to solve the issues presented.

Carefulness and accuracy are needed in the preparation of questions (Yunita et al., 2019). The statement which stated that it cannot use the items in the form of multiple-choice to assess HOTS validly was not valid. Even in some instances (Sani, 2016), written tests in the form of multiple choices are also more efficient if used to determine the essential competencies of students. However, one must admit that multiple-choice tests cannot be used to measure creative thinking skills. Likewise, the best instrument for assessing HOTS is a question in the form of a description or open-ended.

The trials conducted on 41 respondents (students who have programmed the Field Analytical Geometry Course) obtained unsatisfactory results. If traced, several factors may be the cause. Among the factors, one of them was because the students tested were students who had passed the flat-plane analytic geometry course. This might cause the inability to remember the materials that have been taught in such a long term. The students' long term memory apparently influences the success of the assessment (Dehn, 2010). It was about three months had passed, the students had already forgotten the concepts and procedures related to the study. As a conclusion, testing is better conducted soon after the lessons.

4. CONCLUSION

Items with flat-plane analytic geometry material developed based on Tessmer's (1990) model development have met the characteristics of higher-order thinking. The attributes in question are as stated in Bloom's Taxonomy. Each item meets the validity of content (content), construct and language. Content validity is the compatibility between the questions and the indicators that have been set for each material. Fulfilling construct validity refers because the instrument measures the HOTS concept. It meets the validity of the language because the sentence structure in the question follows the correct Indonesian language rules and does not contain double meanings or ambiguity. Therefore, these items are appropriate to be used as alternative teaching material for students to develop higher-order thinking skills. Future researchers are suggested to test these items on a larger scale and develop those items for different subjects.

REFERENCES

Chun, T. C., & Abdullah, M. N. L. Y. B. (2019). The teaching of higher order thinking skills (HOTS) in Malaysian schools: Policy and practices. *MOJEM: Malaysian Online Journal of Educational Management*, *7*(3), 1-18.

Dehn, M. J. (2010). *Long-term memory problems in children and adolescents: Assessment, intervention, and effective instruction*. John Wiley & Sons.

Hinde, ER & Perry, N. (2007). Elementary teachers' application of Jean Piaget's theories of cognitive development during social studies curriculum debates in Arizona. *The Elementary School Journal 108*(1): 63-79.

Höst, M., Regnell, B., & Wohlin, C. (2000). Using students as subjects—a comparative study of students and professionals in lead-time impact assessment. *Empirical Software Engineering*, *5*(3), 201-214.

Marshall, JC, & Horton, RM (2011). The Relationship of Teacher Facilitated Inquiry-based Instruction To Student Higher-Order Thinking. *School Science and Mathematics 111*(3): 93-101.

Misri, M. A., Saifuddin, S., Akbar, R. O., & Kamelia, N. R. (2021). Development and evaluation of a HOTS-based test for matrix topic: A classical test and item response theory. *Beta: Jurnal Tadris Matematika*, *14*(2), 126-141.

Nugroho, R. Arifin. 2018. *HOTSSkills Thinking Higher Order: Concepts, Learning, Assessment and Questions*. Jakarta: Grasindo.

Ormrod, Jeanne Ellis. 1995. *Educational Psychology: Principles and Applications*, New Jersey: Prentice-Hall.

Panggabean, Ellis Mardiana, Haryati, Feri & Sri Wahyuni. 2020. Assessment Practices In Learning Mathematics. *International Journal of Mechanical and Production Engineering Research and Development (IJMPERD).* Vol. 10, Special Issue, Aug 2020, 205– 215© TJPRC Pvt. Ltd.

Pratama, HS & Retnawati, H. (2018). *Urgency of Higher Order Thinking Skills (HOTS) Content Analysis in Mathematics Textbook*. Journal of Physics: Conference Series.

Resnick, LB 1992. *Educational and Learning to Think* (Washington DC: National Academy Press).

Sani, Ridwan Abdullah. 2016. *Authentic Assessment*. Jakarta: Earth Literacy.

Santos-Trigo, L. & Moreno-Armella, L. (2013). International perspectives on problem-solving research in mathematics education, a particular issue. *The Mathematics Enthusiast 10*(1&2): 3–8.

Saputro, S., Perdana, R., Atmojo, I. R. W., & Nugraha, D. A. (2020, April). Development of science learning model towards Society 5.0: A conceptual model. In *Journal of Physics: Conference Series* (Vol. 1511, No. 1, p. 012124). IOP Publishing.

Ramadani, M., Supahar, S., & Rosana, D. (2017). Validity of evaluation instrument on the implementation of performance assessment to measure science process skills. *Jurnal Inovasi Pendidikan IPA*, *3*(2), 180-188.

Tajudin, Nor'ain & Chinnappan, Mohan. 2016. *The Link Between Higher Order Thinking Skills, Representation and Concepts in Enhancing TIMSS Tasks*. International Journal of Instruction vol 9 No 2, July 2016.

Tarman, B., & Kuran, B. (2015). Examination of the cognitive level of questions in social studies textbooks and the views of teachers based on bloom taxonomy. *Educational Sciences: Theory & Practice*, *15*(1).

Tessmer, M., & Wedman, J. F. (1990). A layers-of-necessity instructional development model. *Educational Technology Research and Development*, *38*(2), 77-85.

Wagering. 2011. *Classroom Assessment: An Integral Part of the Learning Process to Prepare Resources HumanHolistically*. Available in the journal Invotec, Vol VII, No. 2 August 2011. ISSN 1411-5514 pp. 199-217

Widana, I. W. (2020, July). The effect of digital literacy on the ability of teachers to develop HOTS-based assessment. In *Journal of Physics: Conference Series* (Vol.1503, No. 1, p. 012045). IOP Publishing.

Yunita, D. R., Maharani, A., & Sulaiman, H. (2019, April). Identifying rigorous mathematical thinking on olympic students in solving non-routine problems on geometry topics. In *3rd Asian Education Symposium (AES 2018)* (pp. 495-499). Atlantis Press.