

Analysis of higher order thinking skills (HOTS) mathematics questions

Fitri Alfarisa ^{a,1,*}, Supriadi ^{a,2}, Susilawati ^{a,3}, Annisa Dinur Rahimah ^{a,4}, Rizi Suci Litundzira ^{a,5}

^a Universitas Pendidikan Indonesia, Indonesia

¹ alfarisa@upi.edu*; ² supriadi.upiserang@upi.edu; ³ susilawati@upi.edu; ⁴ annisadinur@upi.edu; ⁵ rizzysuci018@gmail.com

* Korespondensi Penulis

ARTICLE INFO

Article history

Received December 12, 2022

Revised December 14, 2022

Accepted December 15, 2022

Available Online December 16, 2022

Keywords

HOTs

Mathematic

Rasch Model

Primary School

ABSTRACT

This study aims to determine the quality of HOTS mathematics questions for fourth grade elementary school students that have been developed. This research is descriptive research with a quantitative approach. The HOTS mathematical instrument was developed with 15 reasoned multiple-choice questions by considering three aspects of HOTS according to Bloom's Taxonomy namely Analyzing (C4), Evaluating (C5), and Creating (C6). The questions were made using fractional material and tested on 15 grade IV students at SD Laboratorium UPI Kampus Serang. The validity of the HOTS instruments was verified by 3 experts in the field of mathematics and continued with calculations using the Aiken V index formula. The quality of the HOTS math questions developed was seen from the validity, reliability, and item difficulty index. The validation results show that all 15 items are valid with an average Aiken index of 0.94. The reliability of the HOTS instrument is also quite reliable with a reliability index of 0.74 which is calculated using Cronbach Alpha. The index of difficulty level of HOTS items is in the range of -1.03 to 1.04. This indicates that all questions are included in HOTS.

This is an open access article under the [CC-BY-SA](#) license.



1. Introduction

The quality of education is inseparable from the learning process carried out in schools and cannot be separated from evaluation procedures in the form of measuring learning outcomes. This happens because measurement is one of the many factors in the system that determine the success of education. Improving the quality of education can be done by creating a good measurement system (Brookhart & Mcmillan, 2019; Mardapi, 2012). Procedures for measurement activities must be systematic, their implementation must also have high accountability, and the expected results can explain the actual abilities of students. Taking measurements, an educator will easily assess the level of understanding and mastery of students towards the subject matter, even information can easily be collected to what extent students are able to apply, synthesize, and analyze the material taught by educators.

One of the abilities that teachers must have been to make learning evaluations that are able to represent the abilities of their students. One evaluation tool that is widely used by teachers is the test. Istiyono, Mardapi, & Suparno (2014) revealed that the test is a technique used in carrying out measurement activities, in which there are various questions, or a series of tasks that must be done or answered by students to measure aspects of student behavior.

Based on the results of observations made by researchers, math test questions in elementary schools tend to emphasize low-level abilities in learning. Students absorb information passively. This learning does not make students gain experience to develop higher-order thinking skills, especially skills in giving simple explanations and concluding (Abdullah, 2017). Higher Order Thinking Skills or high-level thinking is thinking at a higher level which is not just remembering facts or retelling something heard to others (Retnawati, Djidu, Kartianom, Apino, & Anazifa, 2018). Students will tend to use logic rather than just remembering and memorizing, so students will master concepts and be able to solve a more complex problem.

Higher Order Thinking Skills (HOTS) is a process that must be carried out by students with the aim of processing the information and ideas they have, so that students gain new understanding (Budiman & Jailani, 2014). This is commensurate with what was stated by Preus (in Royantoro, et al., 2018: 373) that HOTS is the ability of students to think and connect the concepts being studied with concepts that have never been studied before.

Higher-order thinking skills are abilities that are essential to life, to work, and to function effectively in all other aspects of life (Apino & Retnawati, 2016). Critical thinking skills are divided into 5 groups according to skills of giving simple explanations, building basic skills, concluding, giving further explanations, organizing strategies and tactics (Pratama & Retnawati, 2018).

Kane, Mishra, & Dutta (2016) suggests the characteristics of HOTS questions, namely: a) measuring students' ability to think at a high level; b) based on problems that are in accordance with the real context; and c) question texts can use various forms of questions in the form of multiple choices, abbreviations, short answers, and descriptions. In addition, Ramos, Dolipas, & Villamor (2013) state that questions that include Higher Order Thinking Skills have the following characteristics: a) move one concept to another; b) processing and implementing data; c) explore a variety of different sources; d) use data (information) to solve problems; and e) examine ideas and data in depth.

Assessment can be implemented to assist students in improving their higher order thinking skills. This is supported by another opinion, that questions of higher order thinking can encourage students to think deeply about the subject matter (Brookhart & Mcmillan, 2019). This indicates that tests of higher order thinking skills can provide stimulation to students to develop their abilities. This is based on the argument that the effectiveness of learning in the classroom cannot be separated from measurement and assessment activities whose information can be used to improve the quality of learning. It is necessary to know the procedure for developing HOTS questions and analyzing the quality of HOTS questions.

2. Method

This study used a quantitative approach with quantitative descriptive methods (Creswell & Creswell, 2017). The instrument developed in this study was 15 math HOTS questions. HOTS math questions are made in the form of reasoned multiple choice with fractional material. HOTS questions were developed from Bloom's Taxonomy theory which contains three aspects, namely analyzing (C4),

evaluating (C5), and creating (C6) (Krathwohl, 2002). The test instruments that have been made are then validated by experts to determine the quality of the tests that have been developed. Validation was carried out by 3 mathematicians (lecturers and teachers/practitioners). Input or expert validation results were collected by means of a questionnaire and then analyzed using the Aiken V index. The reliability of the test instrument was estimated using the Alpha Cronbach formula. The test instrument was tested on 15 grade IV students at UPI Laboratory Elementary School, Serang Campus. The test results were then analyzed to see the level of difficulty of the items and the accuracy of the HOTS aspects which were developed with the help of the WINSTEP program (Boone, Staver, & Yale, 2013; Sumintono & Widhiarso, 2015).

3. Results and Discussion

3.1. Development of HOTS Mathematics Question Items

The math test for fraction material intended for grade 4 elementary school students is made on a HOTS basis. The HOTS aspects used in developing this problem are based on Bloom's Taxonomy, namely C4 (Analyzing), C5 (Evaluating), and C6 (Creating). This HOTS question was created with the aim of measuring students' critical thinking skills.

HOTS math questions are arranged in multiple choice form which consists of one question sentence, called a stem, and several answer choices called alternatives or options. Multiple choice questions have the advantage as stated by Reynolds & Willson (2010), namely they can provide objective and reliable results, are versatile or can be used for everyone and for all theoretical material, can easily analyze the results, and can provide information. diagnosis.

Making HOTS questions requires a modified form of multiple-choice instrument, hereinafter referred to as reasoned multiple choice. This is done to train students to think critically. The development of reasoned multiple-choice questions will increase the variety of items that can be used in assessment, so that the assessment instruments obtained can accommodate broad thinking skills. Questions that contain questions with a high level of thinking will measure the competence of students much more broadly, besides that these questions will require critical thinking skills to solve them.

This HOTS math problem was developed from fraction material derived from 4 competencies, namely: explaining equivalent fractions with concrete pictures and models; explain various forms of fractions (ordinary, mixed, decimal, and percent); identify equivalent pieces with concrete drawings and models; identify the various forms of fractions (regular, mixed, decimal, and percent) and the relationships between them. Furthermore, based on the analysis of the selected basic competencies, it is derived in the HOTS math question grid in Table 1.

Table 1. The HOTS Math Question

Aspect	Sub Aspect	Indicator	No Item
Analyze (C4)	Analyze	Analyze two fractions that are worth	1
	Analyze	Analyze the sequence of fractions from smallest to largest	13
	Analyze	Analyze ordinary fractions in addition and subtraction arithmetic operations	12
Evaluate (C5)	Compare	Compare fractional values	3
	Connect	Connect problems related to equivalent fractions in everyday life.	2
Creating (C6)	Changing	Changing common fractions into mixed fractions and vice versa	4

Aspect	Sub Aspect	Indicator	No Item
	Changing	Changing common fractions into decimal form and vice versa	5, 6, 14
	Changing	Changing ordinary fractions into percentages and vice versa	7, 8, 15

After the grid is made then the items are developed based on the grid. The items that have been compiled are then validated in terms of content to see the suitability of the indicators with the items that have been developed. The math HOTS questions were validated by 3 experts consisting of 2 lecturers and 1 teacher/practitioner. Experts or mathematicians validate the instrument with a questionnaire and proceed by using the Aiken V index calculation. Based on the analysis in table 2, all items are said to be valid. According to the Aiken table (1980) with the number of questions, namely 15 items and using 4 scales, the minimum item is said to be valid, which is 0.73. The overall item mean is 0.94. This indicates that the 15 HOTS math items are classified as valid, good and ready to be tested on a small scale. The results of calculating the Aiken index are presented in Table 2.

Table 2. The HOTS Math Question

Item	E1	E2	E3	S1	S2	S3	ΣS	V
1	4	4	4	3	3	3	9	1.00
2	4	3	4	3	2	3	8	0.89
3	3	4	4	2	3	3	8	0.89
4	4	3	4	3	2	3	8	0.89
5	3	4	4	2	3	3	8	0.89
6	4	4	4	3	3	3	9	1.00
7	3	4	4	2	3	3	8	0.89
8	3	4	4	2	3	3	8	0.89
9	4	4	4	3	3	3	9	1.00
10	4	4	4	3	3	3	9	1.00
11	4	4	4	3	3	3	9	1.00
12	4	4	4	3	3	3	9	1.00
13	4	4	4	3	3	3	9	1.00
14	4	3	4	3	2	3	8	0.89
15	4	3	4	3	2	3	8	0.89

After the content validation was carried out, HOTS math questions were tested on 15 grade IV elementary school students at UPI Serang Laboratory Elementary School. The test results were then analyzed to determine the quality of the HOTS items. Based on the test results of the HOTS mathematics instrument, it is known that the estimated reliability of this instrument is 0.748, which means that this instrument is reliable. These results show that the reliability of the HOTS math instrument is good with a reliability of more than 0.70 (Barnard-Brak, Lan, & Yang, 2018).

3.2. Analysis of Math Hots Items

The results of the analysis of HOTS math questions with the help of the WINSTEP program provide information that the item difficulty index lies in the range -1.03 to 1.04. Items 10 & 11 are items that have the highest level of difficulty with an index value of difficulty level of 1.04, while item 7 is an item that has the lowest level of difficulty, namely -1.03. The characteristics of the item parameters described are in the form of item difficulty levels from -2 to 2. This can be interpreted that the 15 items meet the criteria for a good level of difficulty. Details can be seen in Table 3.

Table 3. Analysis of Item Difficulty Level

Item	Difficulty	Item	Information
10	1,04		Difficult
11	1,04		Difficult
4	0,37		Keep
3	0,21		Keep
5	0,21		Keep
12	0,16		Keep
13	0,05		Keep
14	0,05		Keep
9	-0,11		Keep
15	-0,17		Easy
1	-0,23		Easy
8	-0,41		Easy
2	-0,48		Easy
6	-0,71		Easy
7	-1,03		Easy

The difficulty level of the item can also be seen from the percentage of respondents who answered correctly for each category. If the percentage of respondents who can answer category 1 is the most, it means that the item is difficult to describe, and vice versa, if the percentage of respondents is the most in category 4, it means that the item is easy (Bond & Fox, 2013; Sumintono & Widhiarso, 2014). The percentage of respondents who correctly answered the items in each instrument aspect and sub-aspect for the four categories is stated in Table 4.

Table 4. The Percentage of Answers Based on Aspects and Sub-Aspects

Aspect	Sub Aspect	Category (%)			
		1	2	3	4
Analyze (C4)	Analyze	44,3	2,6	2,6	50,5
Evaluate (C5)	Compare	25	0	0	75
	Connect	58	0	0	42
Create (C6)	Change	36,8	1,10	1,1	51
	Solving	75,3	0	0	24,6

Based on the results in Table 10, it is known that the aspect of analyzing is the easiest aspect, namely 50.5% of students can do well. While the aspect of creating (C6) has the highest level of difficulty, namely only 24.6% of students can work optimally on the finishing sub-aspect. This is in line with Bloom's taxonomy that the aspect of creating is the highest realm of thinking (Krathwohl, 2002). An example of HOTS math questions developed on the aspect of creating or creating is presented in Table 5.

Table 5. Examples of HOTS Questions

Aspects	Creating/Completing (C6)
Basic Competencies	4.2 Identify various fractional forms (regular, mixed, decimal, and percent) and relationships among others
Indicator	Solving problems related to ordinary and decimal fractions in everyday life
Steam	Meli accompanies Mom to shop at the traditional market. Mom bought 5 kg of rice, 2.5 kg of chicken meat, 1/2 kg of carrots, and 3/4 kg of potatoes. Mom's whole grocery is.... A. 7.5 kg

	B. 8.75 kg C. 9.5 kg D. 9.25 kg Reason: A. 8.75 kg is mom's total shopping, including 5 kg of rice, 2.5 kg of chicken meat, kg of carrots, and $\text{kg}\frac{1}{2}\frac{3}{4}$ of potatoes B. 7.5 kg was obtained from the overall results of mom's shopping, which included 5 kg of rice, 2.5 kg of chicken meat, kg carrots, and kg of potatoes $\frac{1}{2}\frac{3}{4}$ C. 9.5 kg was obtained from the overall results of mom's shopping, which included 5 kg of rice, 2.5 kg of chicken meat, kg carrots, and kg of potatoes $\frac{1}{2}\frac{3}{4}$ D. 9.25 kg is the total of mom's shopping, including 5 kg of rice, 2.5 kg of chicken meat, kg of carrots, and $\text{kg}\frac{1}{2}\frac{3}{4}$ of potatoes			
Answer and Scoring Keys	CATEGORY	ANSWER		SCORE
		QUESTION	REASON	
	Learners answer WRONG questions, and reasons are also WRONG	×	×	1
	Learners answer questions CORRECTLY , and reasons WRONG	B	×	2
	Learners answer WRONG questions, and CORRECTLY reasons	×	A	3
	The learner answers the question CORRECTLY , and the reason is also CORRECTLY	B	A	4

In the question above, the HOTS aspect used is creating (C6). According to (Tan & Halili, 2015) the creating stage directs students to connect elements together and become a unified whole. In this question, students are expected to be able to solve problems related to ordinary fractions and decimals in everyday life. In word problems, students are stimulated to think critically and creatively to be able to interpret concrete problems into abstract forms (Tanujaya, Mumu, & Margono, 2017). The material presented is about the addition of fractions. In common fractions, students can add up the numerator and numerator with the same denominator. If the denominators are different, they must be equated first. However, in the questions above, students are faced with different forms of fractions, namely ordinary fractions, and decimal fractions. Therefore, students must choose to add up in the form of ordinary fractions or in the form of decimal fractions.

4. Conclusion

The development of HOTS mathematics questions that contain fractional material for class IV Elementary School is carried out with the aim of measuring students' critical thinking skills. The HOTS item instrument was developed based on Bloom's Taxonomy including Analyzing (C4), Evaluating (C5), and Creating (C5) Aspects. The HOTS mathematics instrument developed in the form of reasoned multiple choice and has been proven with content validity by 3 mathematicians. The validation results prove that the 15 items developed are valid with an average Aiken index of 0.94. Furthermore, HOTS mathematics questions have a difficulty index that all meet the range of -1.03 to 1.04.

Reference

- Abdullah, A. H. (2017). Mathematics Teachers' Level of Knowledge and Practice on the Implementation of Higher-Order Thinking Skills (HOTS). *EURASIA Journal of Mathematics, Science and Technology Education*, 13(1), 3–17. <https://doi.org/10.12973/eurasia.2017.00601a>
- Apino, E., & Retnawati, H. (2016). Creative Problem Solving to Improve Students' Higher Order Thinking Skills in Mathematics Instructions Thinking. *Proceeding Of 3rd International Conference On Research, Implementation And Education Of Mathematics and Science*, (May), 16–17.
- Barnard-Brak, L., Lan, W. Y., & Yang, Z. (2018). Differences in mathematics achievement according to opportunity to learn: A 4pL item response theory examination. *Studies in Educational Evaluation*, 56, 1–7.
- Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model: Fundamental measurement in the human sciences*. Psychology Press.
- Boone, W. J., Staver, J. R., & Yale, M. S. (2013). *Rasch analysis in the human sciences*. Springer.
- Brookhart, S. M., & Mcmillan, J. H. (2019). Classroom Assessment and Educational Measurement. In *Classroom Assessment and Educational Measurement*. Taylor & Francis. <https://doi.org/10.4324/9780429507533>
- Budiman, A., & Jailani. (2014). Pengembangan instrumen asesmen higher order thinking skill (HOTS) pada mata pelajaran Matematika SMP kelas VIII semester 1. *Jurnal Riset Pendidikan Matematika*, 1(2), 139–151.
- Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.
- Istiyono, E., Mardapi, D., & Suparno, S. (2014). Pengembangan tes kemampuan berpikir tingkat tinggi fisika (pysthots) peserta didik SMA. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 18(1), 1–12.
- Kane, S. N., Mishra, A., & Dutta, A. K. (2016). Developing Instructional Design to Improve Mathematical Higher Order Thinking Skills of Students. *Journal of Physics: Conference Series*, 755(1), 011001. <https://doi.org/10.1088/1742-6596/755/1/011001>
- Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, 41(4), 212–218. <https://doi.org/10.1207/s15430421tip4104>
- Mardapi, D. (2012). *Pengukuran penilaian dan evaluasi pendidikan*. Yogyakarta: Nuha Medika.
- Pratama, G. S., & Retnawati, H. (2018). Urgency of Higher Order Thinking Skills (HOTS) Content Analysis in Mathematics Textbook. *Journal of Physics: Conference Series*, 1097(1), 012147. <https://doi.org/10.1088/1742-6596/1097/1/012147>
- Ramos, J. L. S., Dolipas, B. B., & Villamor, B. B. (2013). Higher order thinking skills and academic performance in physics of college students: A regression analysis. *International Journal of Innovative Interdisciplinary Research*, 4(48–60).
- Retnawati, H., Djidu, H., Kartianom, Apino, E., & Anazifa, R. D. (2018). Teachers' knowledge about higher-order thinking skills and its learning strategy. *Problems of Education in the 21st Century*, 76(2).
- Sumintono, B., & Widhiarso, W. (2014). *Aplikasi model Rasch untuk penelitian ilmu-ilmu sosial (edisi revisi)*. Trim Komunikata Publishing House.
- Sumintono, B., & Widhiarso, W. (2015). Aplikasi Permodelan Rasch Pada Assessment Pendidikan. In *Aplikasi Permodelan Rasch Pada Assesment Pendidikan*.

- Tan, S. Y., & Halili, S. H. (2015). Effective Teaching of Higher-Order Thinking (HOT) in Education. *The Online Journal of Distance Education and E-Learning*, 3(2), 41–47.
- Tanujaya, B., Mumu, J., & Margono, G. (2017). The Relationship between Higher Order Thinking Skills and Academic Performance of Student in Mathematics Instruction. *International Education Studies*, 10(11), 78. <https://doi.org/10.5539/ies.v10n11p78>