

Developing higher-order thinking skill (HOTS) test instrument using Lombok local cultures as contexts for junior secondary school mathematics

*¹Syukrul Hamdi; ²Iin Aulia Suganda; ³Nila Hayati

^{1,2,3}Universitas Hamzanwadi

Jl. Cut Nyak Dien No.85, Pancor, Selong, Lombok Timur, West Nusa Tenggara 83611, Indonesia

*Corresponding Author. E-mail: syukrulhamdi@hamzanwadi.ac.id

Submitted: 23 November 2018 | Revised: 04 December 2018 | Accepted: 10 December 2018

Abstract

The study was aimed at producing a valid and reliable higher-order thinking skill (HOTS) test instrument using Lombok local cultures as contexts in the junior secondary school mathematics subject matter. The study is developmental research involving a field try-out of 75 students of Grade VIII. Data were analyzed using classical test theories of difficulty levels, discriminating powers, and functioning distractors. The test validity is assessed using the Aiken formula and reliability is estimated by Cronbach Alpha. Findings show that, of the 20 initial multiple-choice items, 15 were valid and reliable and had the characteristics of good test items with a medium-rated difficulty level average of 0.28, a good-rated discriminating power of 0.31, a good-rated reliability coefficient of 0.79, and all distractors well-functioning.

Keywords: *test item development, higher-order thinking skill (HOTS), junior secondary school mathematics education*

Introduction

Twenty-first-century education does not merely provide access to information for students. It is expected to form generations to be able to act effectively in facing the complex and ever changing world's challenges. It must be able to give new experiences, unique and creative ideas, and develop collaborative attitudes as learners' capital to face the world of work, get along with society, and live the daily lives. *The Partnership for 21st Century Skill* (Warisdiono, et al., 2017, p. 18) explains that learning in educational world must focus in developing the 4C's as competencies which must be acquired to face the 21st century: creativity, critical thinking, communication, collaboration. This has had a great influence on the educational curricula in accommodating 21st-century competencies into the school subject matters, including mathematics.

Mathematics is one of the knowledge fields that have central roles in the development of competencies needed to face the 21st century environments. Mathematics understanding is a readiness centre for the young generation to live in modern society. A proportion of the growth of problems and situations exposed in daily lives, including in the professional contexts, needs a number of levels of mathematics understanding, mathematics thinking, and mathematics tools. Mathematics is an important tool for the young adults in confronting the issues and problems in the personal, professional, societal, and scientific environments in their daily lives (OECD, 2013 in Kurniati, Harimukti, & Jamil, 2016, p. 143). However, the low level of the learners' mathematics knowledge has attracted the attention of the educators and researchers and has always become hot topics of discussions in society.

A number of international evaluations on mathematics learning reveal that Indonesian students have not shown pleasing reality. Indonesia has 1,095 class hours per year but students' competencies are under the average level, as compared to South Korea that has 903 and Japan with 712, sitting in the high level of the world ranking (Rahmawati, 2016, p. 6). The Indonesian's involvement in the international assessment is how its educational achievement is among other countries in the world. Results of the study of Programme for International Student Assessment (PISA) conducted by Organization for Economic Cooperation and Development (OECD), looking at the thinking abilities of students around the 15 years of age in reading, mathematics, and science since 2000, show that the average score of mathematics literacy of Indonesian children is still under the international standard (Indonesia PISA Center, 2013). The mathematics literacy of Indonesian children is, therefore, low. In the PISA study 2015 that took 540,000 15-year old students from 72 countries, Indonesia was at the 63 rank of the 70 countries being assessed with a mathematics score of 386. The international standard score was 490 (OECD, 2016). This shows that the mathematics literacy average score of Indonesian students is still under the international standard score.

Beside PISA, results from another study, Trend in International Mathematics and Science Study (TIMSS) taken by Indonesia since 1999, reported the same thing. The mathematics competences of grade VIII Indonesian students were low (Scientific Literacy, October 24, 2014). The TIMSS study in 2015 showed that Indonesian students scored 397 out of the international standard 500. Indonesia is still under the average rank, 45 out of 50 countries (Mullis, Martin, Foy, & Arora, 2015). Details of the PISA and TIMSS mathematics ranking of Indonesian students can be seen in Table 1.

This condition of Indonesian education in mathematics is frightening. The PISA and TIMSS studies pointed out that the students lacked logic and reasoning in completing test items that demand the competences of analysis, evaluation, and creation.

Table 1. Mathematics ranking of Indonesian students by PISA and TIMSS

Year	PISA	PISA Score	TIMSS	TIMSS Score
1999 / 2000	39 of 41	367	34 of 38	403
2003	38 of 40	360	35 of 46	411
2006 / 2007	50 of 57	391	36 of 49	397
2009	61 of 65	371	-	-
2011 / 2012	64 of 64	375	38 of 42	386
2015	63 of 70	386	45 of 50	397

(Sources: Indonesia PISA Center, 2013; Mullis, Martin, Foy, & Arora, 2012; Mullis et al., 2015; OECD, 2014, 2016; Scientific Literacy, 2014)

The Director of the National Educational Evaluation Centre (NEEC), Nizam (Krisiandi, 2016), stated that Indonesian students are good at answering questions of the memorization type, but poor at application and reasoning. School learning, from daily quizzes to school exams, has not sharpened students' abilities to reason. Nizam also mentioned that learning through the subject matters must not be directed only to knowledge skills but also to competences. In the 21st century, basic literacy (science, mathematics, reading, and technology) and also competencies of critical, creative, communicative, and collaborative thinking must be mastered.

The NEEC researcher, Rahmawati (Krisiandi, 2016), also stated that the students' competences in higher-order thinking are still weak; students must be habituated with higher-order thinking test items. Teachers are expected to develop test items which deal with higher-order thinking. This is not as easy; yet, teachers need to familiarize themselves with high-order test items, items which are used by TIMSS and PISA. This is in agreement with the National Curriculum 2013 that demands learner competencies to communicate and think critically and creatively. The study by Kurniati et al. (2016, p. 154) had the same tone with the NEEC study by Rahmawati stating that the lack of higher-order thinking skills (HOTS) in the students is caused by the inability of the students to understand the subject-matter material and apply it in daily life.

Revision to the Curriculum 2013 in 2017 requires teachers to make a number of improvements. Among others, one is for the teacher to be creative in integrating literacy, 21st 4C skills (creative, critical, communicative, and collaborative), and HOTS in their classroom instruction (Pedia Pendidikan, 2017). Phol (Kurniati et al., 2016, p. 143) stated that the ability to involve analysis, evaluation, and creativity is a higher-order ability. According to Brookhart (2010, p. 29), the HOTS involves logic and reasoning, analysis, evaluation, creation, problem solving, and also judgment. Further, Hamdi, Kartowagiran, and Haryanto (2018, p. 1) stated that, at the third level, which is high level, students' understanding is characterized by the abilities to work with complex materials such as mathematical thinking and reasoning and communicative, critical, creative, interpretative, reflective, generalizing, and mathematical skills.

The use of HOTS items in tests is able to train students to sharpen their abilities and skills that are in line with the 21st-century demands. Through HOTS-based test items, critical thinking skills (creative thinking and doing, creativity, and self-reliance learning), will be built through practices in solving various daily-life real problems (problem-solving) (Warisdiono, et al., 2017, p. 18).

The elevation of higher-order thinking skills has become a priority in the school mathematics learning. Students of the junior secondary levels must be trained toward higher-order thinking in accordance with their age. This can be done by the teacher by giving test items of the HOTS type. For this, it is not enough for the teacher to merely pick up material from the packaged workbooks; but they need to resource to more weighted materials. The problem faced by teachers is that they have insufficient ability to develop test items of the HOTS type.

At school, many teachers still use test items that tend to test students' memory aspects rather than higher-order thinking skills. The test items are directed more to lower-thinking skills (LOTS) of memorization and understanding. On the other hands, what the students need to face the future demands is HOTS. The development of HOTS in stu-

dents is expected to raise students' ability in problem solving, elevate their self-confidence in mathematics, and improve their learning achievement (Butkowski, et al., 1994 in Budiman & Jailani, 2014, p. 142).

A HOTS test item is given through a stimulus. A stimulus can be derived from the recent global issues such as technology, information, science, education, health, and infrastructure. A stimulus can also be raised from the environment such as cultures. It is a fact, however, that test items in the school books lack the involvement of cultural issues. In fact, peoples like the Japanese, Chinese, Koreans, and others have used cultural issues in their mathematics learning which makes far advanced in all fields. Kurumeh stated that the success of the Japanese and Chinese in mathematics learning is because they use ethnomathematics (Supriadi, Arisetyawan, & Tiurlina, 2016, p. 2).

Various cultural products of the Indonesian ancestors show art creativities that contain mathematics elements. The case is the same with the cultural products of the Lombok Sasak tribes. One example is the shield from Ende used in a traditional dance. It is made of thick buffalo leather with a two-dimensional geometric pattern. Another product is the Sasak house architecture with three-dimensional ornaments. Besides, many traditional clothes of Sasak have geometrical pattern motifs and the traditional wedding ceremonies have statistical elements. One example is presented in Figure 1.



Figure 1. Example of cultural products of Lombok Sasak ethnic (Department of National Education, 2000, p. 21)

According to Acho, Imako, and Uloko (Wulandari & Puspawati, 2016, p. 35), students' memory and learning achievement obtained through cultural-based instruction are higher than those obtained through conventional teaching. Besides, the study by Nur and

Palobo (2017, p. 11) shows that contextual instruction using Lombok local cultures as contexts gives a positive and significant influence on the students' problem-solving abilities in mathematics.

In addition, Curriculum 2013 requires teachers to be able to develop HOTS test items in line with local environments. As such, stimuli for the test items will be attractive since they can be directly observed and accepted by students. Besides, the use of local cultures for HOTS test items will increase students' senses of attachment and ownership towards the local potentials of their place. Linking mathematics with cultures will expectedly help students see the connection and application of mathematics not only with other disciplines of science, but also with real life.

The item format developed in the present study is that of the multiple-choice type. According to the opinions of experts and research results, tests of the multiple-choice format can be used for HOTS (Budiman & Jailani, 2014, p. 142). The procedure suggested for the HOTS items is that of a set of items consisting of an input followed by answer options.

Based on the rationalization added with data and supporting evidence presented, a need is felt on developing HOTS test instruments with local cultural contexts in the mathematics subject matter of the junior secondary school to prepare students to face the 21st century. The valid and reliable test instrument can be used to train students' in attaining HOTS, help teachers in testing students' HOTS, and become a reference source for the development of HOTS test items for other base competencies in the syllabus.

Method

The study was development research. It applied the seven steps of gathering initial information, planning, development of first draft and expert validation, limited-scale try-out/readability, revision of the first draft, field try-out, and revision of the final product.

Initial information gathering was related to the product to be developed. It was done through theoretical reviews covering needs

analyses, reviews of the concepts and theories concerning HOTS and local cultures, and analyses of the core competencies (CC) and base competencies (BC) of the Semester 2 of Grade VIII of junior secondary mathematics in the Curriculum 2013.

In the planning phase, the design of the developed product was outlined through the steps of defining, formulating the objectives, and designing of the initial product. This consisted of formulating the product specification, determining the objectives, and constructing the table of specification for the HOTS test items using Lombok cultures as the contexts.

The developed HOTS test items were constructed using HOTS indicators and also BC indicators. The HOTS indicators were synthesized from Ennis (Komalasari, 2013, p. 266; Bayer, Ellis, Gokhale, Cotton, & Langrehr (Thebooke, n.d.); Torrance (Lestari & Yudhanegara, 2015, p. 89); Budiman & Jailani, 2014, p. 143). The indicators were (1) identifying and relating relevant information from a problem; (2) making accurate conclusion based on the obtained information; (3) finding consistencies/inconsistencies in the product; (4) evaluating the product against determined criteria/standards; (5) synthesizing ideas/strategies for the problem solution; (6) applying the strategy for the problem solving; and (7) developing new alternatives of the problem solving.

In the development of the initial product, a first draft of the HOTS instrument was developed. This consisted of 20 multiple-choice test items.

The draft was then subjected to validation by the expert team from the Department of Mathematics Education. The objective of this assessment was to see whether or not the developed test was acceptable and feasible to be used. Another purpose was to obtain feedback for the improvement of the draft.

After being validated by the experts, the draft was then subjected to the analysis of the results of the item validation. Data were in the form of scores of the test items by the experts. The analysis used Aiken's V formula to calculate the content validity index of the test items.

The next step was trying out the draft in a limited-scale group. The validated and revised draft was tried out in a group of 15 junior secondary school students. The try-out was done to obtain information concerning the testees' ease measure in reading the items, level of attractiveness of the test, and level of testees' interest in the test. The results were converted into percentages wherein $\geq 60\%$ means positive.

The following phase was the revision of the first draft based on the results of the limited-scale try-out. After being revised, the product was then subjected to the field try-out. The field try-out was conducted in two Grade VIII classes in two junior secondary schools. These schools were MTs. Muallimin NW Pancor and MTs. NW Pancor. This try-out involved 75 students. The resulting data were analyzed empirically by way of classical test parameters.

The final step was the revision of the product. This was done on the second draft that was tried out in the two schools. An item was accepted as a final product if it fulfilled one of the following criteria: (1) The item satisfied all the requirements of difficulty levels, discriminating powers, and functioning distractors; and (2) Easy and difficult items were accepted if they had a discriminating power of the good/medium category and the placement of the distractors was functioning. The items that were accepted were then re-formatted to become final products verified as HOTS items.

Findings and Discussion


Findings

Higher-order thinking skills include critical thinking, creative thinking, and problem solving. Problem solving, seen as the main skill in HOTS, is a skill in critically and effectively managing, combining, or developing information in the form of facts or ideas to solve a problem and make a decision or finding a solution to a hard-to-handle situation. A HOTS item is one that requires the ability to apply higher-level thinking. The item is presented using a stimulus. A stimulus can be sourced from global issues such as techno-

logy, information, science, education, health, and infrastructure. A stimulus can also be obtained from the environment such as cultures.

Lombok is one of the islands in Indonesia which retains various cultures from history in the forms of objects, non-objects, traditional habits, ethics, and arts. The diversity of the cultures can be used as stimuli and integrated into the school learning processes, including mathematics. Inheritances of history, architecture, dances, musical instruments, and others contain mathematics elements. One example is the shield from Ende used in a traditional dance. It is made of thick buffalo leather with a two-dimensional geometric pattern. Another product is the Sasak house architecture with three-dimensional ornaments. Besides, many traditional clothes of Sasak have geometrical pattern motifs and the traditional wedding ceremonies have statistical elements. HOTS items can be integrated with cultural elements. One example of the development of test items based on cultural elements, in this case, Lombok, can be seen in Figure 2.

Gendang Belek merupakan musik khas Suku Sasak di Lombok. Beberapa peralatan musik *Gendang Belek* memiliki bentuk yang sama dengan ukuran yang berbeda-beda seperti gambar di bawah ini



Pencsek Kenceng Terumpang

Panjang jari-jari *Kenceng* sama dengan dua kali jari-jari *Pencsek*. Panjang jari-jari *Terumpang* sama dengan tiga kali jari-jari *Pencsek*. Jika L_1 , L_2 , dan L_3 berturut-turut menyatakan luas alat musik tradisional *Pencsek*, luas alat musik tradisional *Kenceng*, dan luas alat musik tradisional *Terumpang*. Tunjukkanlah mana pernyataan yang benar di bawah ini!

I. $L_1 + L_2 > L_3$ II. $L_1 + L_2 < L_3$
III. $L_1 + L_2 = L_3$

Figure 2. Example of HOTS item using the context of Sasak culture

Translation:

Gendang Belek is a music instrument specific to Sasak ethnic in Lombok. Some *gendang belek* have the same form but different sizes, as can be seen in the pictures.

The radius of *Kenceng* is twice as large as that of *Pencek*. The radius of *Terumpang* is the same as thrice as that of *Pencek*. If L1, L2, and L3 consequently states the size of *Pencek*, *Kenceng*, and *Terumpang*, which of the following statements is correct?

Figure 2 shows an example of a HOTS test item using a local cultural context of Lombok with a HOTS indicator of critical thinking (Making an accurate conclusion from the information of a situation/problem). To be able to answer the question, the testee needs to be able to recall and understand factual, conceptual, and procedural material about circles. Then, by doing an analysis of the situation (stimulus), the testee determines the strategy in solving the problem.

Other than the example above, there are other cultural inheritances that can be integrated into mathematics. The use of cultural elements in HOTS test items will be able to elevate students' senses of attachment and ownership towards the local potentials of their place. Linking mathematics with cultures will also help students see the connection and application of mathematics not only with other disciplines of science but also with the real world.

Instrument Development

The product of the developmental study is a valid and reliable HOTS test instrument, using the local cultures of Lombok as a context, consisting of multiple-choice test items for junior secondary school mathematics. The instrument development passes two assessment phases. The first phase is to assess the validity of the instrument, conducted by three experts of mathematics education. The second involves a limited-scale try-out with 15 testees and a field try-out in two schools with 75 testees.

Validation by experts is to look at the contents of the initial product and obtain feedbacks for revising the first draft. In the process, the experts are given the table of the specification of the test, the test items, and the evaluation sheets. Data of the experts' evaluation are subjected to the Aiken's V formula to find the content validity coefficient. The results can be seen in Table 2.

Table 2. Results of experts' validation

Item Number	Aiken's V Coefficient	Criteria
1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20	0.67-1.00	Good to be used
6	0.33	Need revision/deletion

In Table 2, it can be seen that, out of the 20 test items, 19 are feasible for use and one needs revision or deletion. However, there is a number of items which needs to be revised or deleted following the experts' feedbacks. These include the format of the writing, completeness of the stimulus texts, clearer pictures, and suitability with the junior secondary school level.

The results of the readability check in the limited-scale try-out show that the majority of the students give positive responses towards the test, between 75% and 94%. This is strengthened by positive comments written by some students. Sample statements can be seen in Figure 3. Meanwhile, the difficulty levels of the items can be seen in Table 3.

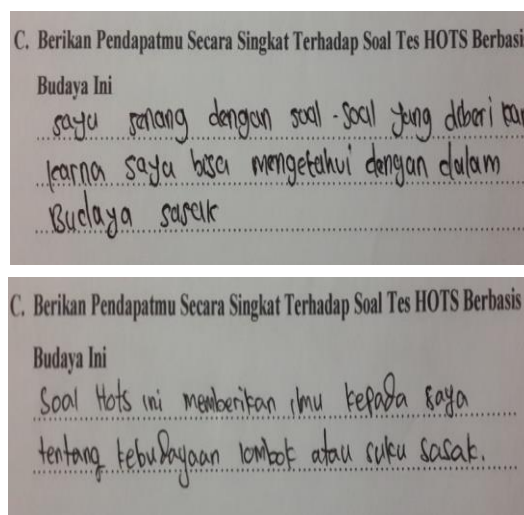


Figure 3. Students' comments on the use of the test instrument

Translation:

Give short opinions about this culture-based HOTS test

I like the test that is given because I can know about Sasak cultures more deeply

This HOTS test give knowledge to me about Lombok cultures or Sasak ethnic

Table 3. Difficulty levels of the main product test items

Category	Item Number	Total
TK < 0,25 (Difficult)	1, 3, 6, 17	4
0,25 ≤ TK ≤ 0,75 (Medium/Enough)	2, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18	14
TK ≥ 075 (Easy)	-	0

Table 3 shows that 14 items (77.78%) have the difficulty level in the medium category. Meanwhile, Table 4 shows that seven test items (38.88%) have a discriminating power of the medium category. The spread of the distractors of the main product test items can be seen in Table 5.

Table 4. Discriminating power of the main product test items

Category	Item Number	Total
DP < 0.20 (Poor)	1, 6, 9, 14	4
0.20 ≤ DP < 0.40 (Medium)	4, 12, 13, 15, 16, 17, 18	7
0.40 ≤ DP < 0.70 (Good)	2, 3, 5, 7, 10, 11	6
0.70 ≤ DP ≤ 1.00 (Very Good)	8	1

Table 5. Effectiveness of distractors of the test items

Category	Item Number	Total
Functioning	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18	18
Not functioning	-	

In Table 5, it is clear that the distractor distribution of all of the test items is functioning; it means that all the distractors are chosen by 5% of the testees. Based on the results of the analyses of the item characteristics above, the number of items that are accepted and replaced/rejected can be seen in Table 6.

In Table 6, a total of 15 items (83.33%) are accepted and 3 (16.67%) are rejected. The accepted are then reformatted to become the final product test instrument of HOTS in terms of the test validity.

Table 6. Results of analyses of item characteristics

Category	Item Number	Total	Percentage
Accepted	2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18	15	83.33%
Rejected	1, 6, 14	3	16.67%


Revision of Final Product

The final product revision is conducted to obtain a test instrument that is valid and reliable. Revision is done by looking at the results of evaluation in the two product try-outs. The revision involves experts' validation, limited-scale try-out, and field try-out.

The experts' validation and product try-outs are used as the main consideration for revision. First, item revision is based on the experts' inputs and suggestions. In general, these include the format of the writing, completeness of the stimulus texts, clearer pictures, and suitability with the junior secondary school level. Figures 4, 5, and 6 show items that are good after revision and items that are rejected.

Good item after revision

Bumbungan is a traditional house of Sasak Ethnic in Lombok. *Bumbungan* has a steep roof, made from hay with a thickness of about 15 cm. The roof is intentionally let to span to the bottom wall and almost covers the wall. Like the picture below.




The roof bottom is 5.2 in length and the top 5/13 of the roof bottom. Height of the roof is 3/2 of the roof top. The circumference of the roof of the house is ...

A. 12.6 m C. 13.6 m
B. 13.4 m D. 14.0 m

Figure 4. Good item after revision

The item in Figure 5 is rejected because the item is not well-formulated and the picture is meaningless (the notes are not clear). Meanwhile, the item in Figure 6 is deleted because it is considered too difficult for junior secondary school age.

Poor item to be deleted



Rudat Dance

The Lombok-specific *Rudat* dance is usually used to welcome guests, involving 10 dancers. The distance of the most-front dancer and the most-back dancer is ... unit.

A. 3.18 C. 6.36
B. 3.56 D. 10

Figure 5. Poor item that is deleted (1)

Poor item to be deleted



Begasingan Lombok Traditional Game

In the picture, an arch is made with the center point of *P* and it crosses the line in *Q* point. Then, with the same radius, an arch is made with the center of *Q*, so that it crosses the first arch in *R* point. From the points of *P*, *Q*, and *R*, *PRQ* angle is made. The size of the angle formed by *PRQ* angle is...

A. 30° C. 60°
B. 45° D. 75°

Figure 6. Poor Items that are deleted (2)

Second, item revision from the limited-scale try-out is done on the results of the analyses of the item characteristics. Most of the revision deals with discriminating powers and non-functioning distractors.

Third, item revision from the field try-out is done in the same way. All the test items are then verified with the HOTS indicators to make sure that all indicators have been represented. After being verified, all items are reformatted to become the final product of the study.

The field try-out involves 75 students, consisting of 24 from MTs. Muallimin NW Pancor and 51 from Mt. NW Pancor. In general, the achievement of students who take parts in the study can be seen in Figure 7.

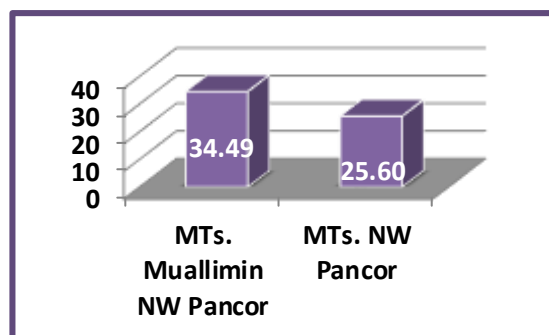


Figure 7. Student achievement profile in mathematics learning viewed from the completed results of the test instrument

Based on Figure 7, it is clear that the average score of students of MTs. Muallimin NW Pancor is higher than that of MTs. NW Pancor. Moreover, a total of 11 students of MTs. Muallimin NW Pancor have scores above the average and 11 have scores below the average. For MTs. NW Pancor, 27 students are above the average score and 24 students are below.

Discussion

The product of the study is a valid and reliable HOTS test instrument using Lombok cultures as contexts. It is a fact that, up to the present time, no effort has been done for evidence of test validity and reliability. The development of the instrument begins with the review of HOTS which, according to Brookhart (2010, p. 29), consist of the ability

of logic and reasoning, analysis, evaluation, creation, problem solving, and also decision making (judgment).

It is followed by formulating the item indicators and writing of the test items. Then, the test items are subjected to content validation through expert judgment. This is followed by the Aiken analyses. Before being administered in the field try-out, the items are subjected to a limited-scale try-out for readability. The field try-out involves 75 students from two schools. Finally, item analyses and reliability estimation are conducted. The test instrument development has been conducted following the standard procedure and found that the test is valid and reliable.

The test items developed in the study are those of the multiple-choice type. According to opinions and research results from experts, a multiple-choice test can be used to measure HOTS (Budiman & Jailani, 2014, p. 142). It is suggested that the format of the HOTS test items consist of an introduction followed by response options.

Conclusion and Suggestions

Based on research findings and discussion, a conclusion is drawn as follows. The final product of the study is a HOTS test instrument using Lombok local cultures as contexts for junior secondary school mathematics consisting of 15 multiple-choice test items with four options. The validity of the test is indicated by the experts' judgment showing that the test is good to be used in the aspects of contents, format, and language. Based on the classical test theories, the instrument fulfills the requirement for reliability shown by a reliability coefficient of 0.79 (good category), with an average score of 0.28 for difficulty levels (medium category), discriminating powers of 0.31 (good category), and functioning distractors.

Based on the conclusion of the study, it is suggested that further research is conducted by analyzing the test items using the IRT as the more modern method. This will expectedly be able to compare the item difficulty levels and the testees' abilities across time and location.

References

- Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. Alexandria: ASCD.
- 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), 139–150. <https://doi.org/10.21831/jrpm.v1i2.2671>
- Department of National Education. (2000). *Kain songket Lombok*. Nusa Tenggara Barat: Kantor Wilayah Provinsi Nusa Tenggara Barat Bagian Proyek Pembinaan Permuseuman.
- Hamdi, S., Kartowagiran, B., & Haryanto, H. (2018). Developing a testlet model for mathematics at elementary level. *International Journal of Instruction*, 11(3), 375–390. <https://doi.org/10.12973/iji.2018.11326a>
- Indonesia PISA Center. (2013). *Ranking Indonesia dalam PISA (2000–2012)*. Retrieved February 11, 2018, from www.Indonesiapisacenter.com/2013/08/ranking-Indonesia-dalam-pisa-2000-2012.html
- Komalasari, K. (2013). *Pembelajaran kontekstual: Konsep dan aplikasi*. Bandung: PT Rafika Aditama.
- Krisiandi. (2016, December 15). Daya imajinasi siswa lemah. *Kompas*, p. 11. Retrieved from <https://nasional.kompas.com/read/2016/12/15/23091361/daya.imajinasi.siswa.lemah>
- Kurniati, D., Harimukti, R., & Jamil, N. A. (2016). Kemampuan berpikir tingkat tinggi siswa SMP di Kabupaten Jember dalam menyelesaikan soal berstandar PISA. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 20(2), 142–155. <https://doi.org/10.21831/pep.v20i2.8058>
- Lestari, K. E., & Yudhanegara, M. R. (2015). *Penelitian pendidikan matematika*. Bandung: PT Rafika Aditama.

- Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international result in mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2015). *TIMSS 2015 international result in mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center.
- Nur, A. S., & Palobo, M. (2017). Pengaruh penerapan pendekatan kontekstual berbasis budaya lokal terhadap kemampuan pemecahan masalah matematika. *AKSIOMA: Jurnal Pendidikan Matematika*, 6(1), 1–14.
- OECD. (2014). *PISA 2012 result in focus: What 15 year olds know and what they can do with what they know*. Paris: OECD Publishing.
- OECD. (2016). *PISA 2015 result in focus*. Paris: OECD Publishing.
- Pedia Pendidikan. (2017). *Penjelasan singkat perbedaan RPP K13 Edisi Revisi 2017 dengan RPP K13 Revisi 2016*. Retrieved February 8, 2018, from <http://www.pediapendidikan.com/2017/05/rpp-k13-revisi-2017.html>
- Rahmawati, S. (2016, December 14). Seminar hasil penilaian pendidikan. *Seminar Hasil TIMSS 2015*. Retrieved from puspen.dik.kemendikbud.go.id/seminar/index.php?folder=hasil_seminar_puspendik202016
- Scientific Literacy. (2014, October 24). *Survei internasional TIMSS (Trends In International Mathematics and Science Study)*. Retrieved February 11, 2018, from literacyofscientific.blogspot.co.id/2014/10/survei-internasioanl-timms-trends-in.html.
- Supriadi, S., Arisetyawan, A., & Tiurlina, T. (2016). Mengintegrasikan pembelajaran matematika berbasis budaya Banten pada pendirian SD Laboratorium UPI Kampus Serang. *Mimbar Sekolah Dasar*, 3(1), 1–18. <https://doi.org/10.17509/mimbar-sd.v3i1.2510>
- Thebooke. (n.d.). *Kemampuan berpikir kritis dan kreatif*. Retrieved March 15, 2018, from <http://thebooke.net/do/download-gratis-buku-berpikir-kritis>
- Warisdiono, et al. (2017). *Modul penyusunan higher order thinking skill (HOTS)*. Jakarta: Direktorat Pembinaan SMA, Direktorat Jenderal Pendidikan Dasar dan Menengah Departemen Pendidikan dan Kebudayaan.
- Wulandari, I. G. A. P. A., & Puspawati, K. R. (2016). Budaya dan implikasinya terhadap pembelajaran matematika yang kreatif. *Jurnal Santiaji Pendidikan*, 6(1), 31–37.