

The Assessment Instrument of Mathematics Learning Outcomes Based on HOTS Toward Two-Dimensional Geometry Topic

I Gede Wahyu Suwela Antara¹, I. Komang Sudarma², I. Ketut Dibia³

^{1,3} Department of Primary Education, Ganesha University of Education, Singaraja

² Department of Education, Psychology, and Guidance, Ganesha University of Education, Singaraja
e-mail: wahyusuwela@gmail.com, ik-sudarma@undiksha.ac.id, ketutdibia31@gmail.com

Abstract

This study aims to developing a mathematics assessment instrument based on Higher Order Thinking Skills (HOTS); and describe the quality of the instrument. This study was a research and development study adapting 4D model from Thiagarajan. The model including the following steps: define, design, develop, and disseminate. Due to limited of time, this research was only carried out until the developing step. The result shows that the instrument that consists of 18 essay test item are valid and appropriate to be used. The instrument reliability coefficients are 0.659 (High). The instrument has the average of item discrimination 0.44 (Very Good) and the average of item difficulty of the instrument are 0.584 (Medium). The conclusion is the assessment instrument is feasible being as an assessment instrument to measure the high order thinking skill toward two-dimensional geometry topic.

Keywords: Assessment Instrument; Mathematics; Higher Order Thinking Skills; Two-Dimensional Geometry; 4D model.

1. Introduction

Higher order thinking skills (HOTS) is defined as the ability to thinking in high level of cognitive domain which includes the ability to analyzing, evaluating, and creating (Irmawati et al., 2018). Higher order thinking skills are very important to be developed in students with regard to 21st century skills demands which include: critical thinking, creative and innovative, communication, collaboration, and confidence (Ariyana et al., 2018). Higher order thinking skills must be owned by students in primary education as an effort to increase the critical and creative human resources in face of competition in the industrial revolution era (Nurhasanah & Yarmi, 2018; Widiastuti & Suyata, 2014).

HOTS based learning is one of the priorities in the mathematics subject in the school. Through the application of HOTS based learning, mathematics learning goals can be achieved optimally (Budiman & Jailani, 2014). Mathematics learning aims to give the students with ability to think logically, analytically, systematically, critically, creatively, and able to make a collaboration (Kemendikbud, 2016). In the primary education, especially elementary school, mathematics aims to make students able to understanding the relationship between mathematical concepts and applying them to solving the daily problems (Japa & Suarjana, 2015).

The low ability of student's high order thinking skills has become a urgent problem in Indonesia (Erfianti et al., 2019). Students often not able applying the knowledge and skills that they have in daily live, not able to resolve problem which slightly differs from they have studied (Kusuma et al., 2017). Along with this, several international assessment result such as PISA and TIMSS also illustrate that the higher order thinking skills of Indonesian students is still below from world average (Tanujaya, 2016).

Based on the results of the analysis of assessment instrument of mathematics for grade IV at SD Negeri 1 Tajun, it was found that the instrument was dominated by questions with C1-C3 cognitive levels. The percentage of cognitive levels in those instruments is 12% for C1, 16% for C2, 56% for C3, 12% for C4, 4% for C5, and 0% for C6. The result of instrument analysis can prove that levels of C1 to C3 dominate with a percentage of 84% while levels C4 to C6 are rarely used with a percentage of only 16%.

The imbalance between the ideal conditions and real conditions causes need for an effort or solution to overcome it. One of solution that can be done is by improving the quality of the assessment instrument that used. Improving the quality of the assessment instrument

can be done through developing a instrument that focused on the HOTS of students (Riadi & Retnawati, 2014). This study aims to developing the assessment instrument of mathematics learning outcomes based on HOTS, and describe the quality of the instrument.

2. Methods

This study was a research and development study. The product that developed in this research was an assessment instrument of mathematics learning outcomes based on HOTS. The instrument that was developed in the form of essay item with cognitive level C4-C6 based on Bloomian Taxonomy (analyzing, evaluating, and creating). This developmental research adapting the 4D (four-D) development model from Thiagarajan. The steps of the model is define, design, develop, and disseminate (Sugiyono, 2011). Due to limited of time, the research was only carried out until the develop step.

The procedures were performed in this research include : identification of problems/needs through observation, interviews, and analyzing assessment instruments used in the schools; prepare a draft of the assessment instrument; conduct an experts judgement to judging in terms of material, language, and construct of the assessment instrument; revise the assessment instrument according to advice from expert judges; instrument trials; tabulating the result; data analyzing; and revise the instruments to produce the final product.

The data in this study include qualitative and quantitative data. Qualitative data were obtained from the results of experts judgement, while the quantitative data were obtained from instrument trials. The results of expert judgement rated based on material, construct, and language to determine the relevance of the test item developed. The result were analyzed by using Gregory formula to determine the content validity.

The instrument trials was tested individually. The subject of instrument trials were all fifth and sixth grade students of SD Negeri 1 Tajun, totaling 70 students. The results of instrument trials were analyzed to determine the items validity, items discrimination, items difficulty, and reliability coefficient of the instrument. Data obtained from instrument trials were analyzed with classical theory parameters and calculated with the help of *Microsoft Office Excel 2016*.

3. Results and Discussion

Expert Judgement (Content Validity Test)

The content validity of the instrument was tested through expert judgement. The expert judgement was carried out by two experts who were leturers of the Department of Primary Education at the Ganesha University of Education. The validation is carried out by providing the draft of instrument with its completeness to the judges. Then, the results of expert judgement was analized by the Gregory formula to calculate the coefficient of content validity. Analysis of the content validity are presented in Table 1.

Table 1. Results of Content Validity

Item Number	Content Validity Coefficient	Criteria
1-20	1,00	Very High

Based on the results of the analysis using the Gregory formula, the draft of instrument consisting of 20 essay items has a content validity coefficient of 1,00. If the coefficient converted into content validity criteria according to Candiasa (2010), the draft of instrument has a very high content validity. However, there were a correction and revision in accordance with the results of the expert judgement, regarding correlation of indicators and items, construct, and language.

Items Validity Test

The items validity test was carried out by field trial at SD Negeri 1 Tajun, totaling 70 students from grade V and VI in the 2019/2020 academic year. The field trial was conducted

on Tuesday, March 17 2020. The instrument that was tested was a instrument that had previously been tested for the content validity through an expert judgement. Total instrument that was tested was 20 essay items. The items validity is calculated by analyzing the results of field trials. The analyzing is conducted by using the product moment correlation formula which is assisted with Microsoft Office Excel 2016. The results of the analysis are presented in Table 2.

Table 2. Results of Analysis of Items Validity

Items Validity	Item Number	Qty	Percentage
Invalid	10, 18	2	10%
Valid	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 19, 20	18	90%

Based on the results of the items validity analysis of the 20 instrument item, there were 18 valid items, and 2 items were invalid. The invalid items were declared null and not used as an instrument.

Item Discrimination Test

Item discrimination tests are conducted on the items that are declared valid (18 items). Item discrimination is calculated using 27% of the total respondents with highest score as a superior group and 27% of the total respondents with lowest score as a asor group. The analysis was assisted with Microsoft Office Excel 2016. The results of the analysis are presented in Table 3.

Table 3. Results of the Analysis of Item Discrimination

Item Discrimination	Item Number	Qty	Percentage
Good	1, 3, 6, 12, 13, 19, 20	7	38.89%
Very Good	2, 4, 5, 7, 8, 9, 11, 14, 15, 16, 17	11	61.11%

Based on the result of the analysis, the average of item discrimination index is 0.44. This value is converted into the criteria of item discrimination according to Koyan (2011) and is in the Very Good criteria.

Item difficulty test

Item difficulty test are conducted on the items that are declared valid (18 items). The analysis was assisted with Microsoft Office Excel 2016. The results of the analysis are presented in Table 4.

Table 4. Results of the Analysis of Item Difficulty

Item Difficulty	Item Number	Qty	Percentage
Easy	1, 2, 3	3	16.67%
Medium	4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 19, 20	15	83.33%

Based on the results of the analysis of item difficulty, the average of item difficulty index is 0.584. This value is converted into the criteria of item discrimination according to Koyan (2011) and is in the Medium criteria.

Reliability Test

The instrument reliability test are conducted on the items that are declared valid (18 items). Instrument reliability was calculated using the Alpha-Cronbach formula. Reliability calculation was assisted with Microsoft Office Excel 2016. Based on the calculation results, the reliability coefficient of the instrument was 0.659. This value is converted into the criteria of item discrimination according to Candiasa (2010) and is in the High reliability criteria.

The results of this research are assessment instrument of mathematics learning outcomes based on higher order thinking skills (HOTS) toward two-dimensional geometry

topic for grade IV in elementary school. This instrument is used to measure students' learning outcomes and train students' higher order thinking skills. The instrument was developed using 4D (four-D) development model. The selection of this model is based on the consideration that this model is very appropriate to be used in developing learning kit, including assessment instruments (Sugiyono, 2011). In addition, the 4D model is also a systematic model, which begins with the analysis of needs in the field so that the instrument that developed depend on of the needs and availability of instruments in the field (Rochmad, 2012).

The coefficient of content validity of the developed instrument are 1.00. This value, after being converted into the content validity category, is in the Very Good category This shows that based on the validation by the validators (judges), the instrument has been theoretically appropriate with the instrumental theory and has a match between the indicators with the items, constructs, and language. Arikunto (2009) argues that an instrument has had suitability with the assessment instrument theory is one of the characteristics of a good instrument and deserves a field trial to determine the empirical quality of the instrument.

The fields trial results show that there were 18 items that were declared valid from 20 items tested. Thow items that were invalid were declared null and not used as instruments. The valid items show that the developed instrument has the appropriate accuracy and precision in carrying out its measurement function (Aritonang R., 2012).

Surapranata (2010) states that instruments with good quality are sown by a coefficient called the coefficient of item discrimination. This coefficient shows the ability of the instrument to distinguish students who have mastered the material and students who have not/do not mastering the material that presented in the instrument. The average of item discrimination coefficient obtained from the analysis of the item is 0.44 with Very Good category. This shows that the instrument has a very ggo ability in distinguishing student who have mastered the material from students who have not.do not master in two dimensional geometry materials.

In general, an instrument is said to have a good quality if the instrument is not too difficult and not to easy (Bagiyono., 2017). The results of the item analysis showed that the average of itme difficulty of developed instruments were 0.584 and were in medium difficultiy level category. Therefore, it can be interpreted that the instruments developed have good quality in terms of item difficulty level.

The reliability coefficient obtained from the analysis of the items was 0.659 with the High category. This shows that the developed instrument product has high testing criteria (reliable) so that it can be tested at any time with fixed or relatively fixed results on equal respondents.

The development results show that the instrument developed has been tested for validity and reliability. This is in line with research conducted by (Budiman & Jailani., 2014) which suggests that the good quality instruments are instruments that have been tested for validity and reliability.

Based on the results of expert validation, fields trials, data analysis, and improvements made, it can be seen that the instrument has fulfilled valid and reliable criteria, and good quality items. So the instruments developed can be used as an assessment instrument on cognitive aspect.

4. Conclusion and Recommendation

Based on results and discussion above, it can be concluded that the final product of this study is the valid and reliable assessment instrument of mathematics learning outcomes based on HOTS toward two-dimensional topic for grade IV students of elementary school. The instrument was development through three development steps, namely defining, designing, and developing. The assessment instrument consists of 18 essay items that have been declared to have an appropriate item validity and content validity. The developed instrument has a reliability coefficient of 0.659 (high). The developed instrument also has an average level of item difficulty of 0.584 (medium) and an average of item discrimination of 0.44 (very good). So that it can be seen that the instruments developed otherwise valid and

has reliability and quality of the items were good. Therefore, the develop instrument can be used as an assessment tools or an accurate learning assessment on cognitive aspect. Some recommendation that can be conveyed based on the results of this study in order to develop learning assessment instruments include: the teacher, if possible, can use an instrument of assessment of high-level thinking skills in mathematics as an evaluation tool on the topic of flat wake in class IV. The teacher should also develop instruments of higher order mathematical thinking ability on other topics. The Principal should facilitate the teacher to develop similar learning instruments so that they can make a positive contribution to the progress of the school and improve the quality of the learning process and outcomes. To other researchers, the results of this study should be used as a reference on learning problems, but by continuing to explore more diverse sources for consideration in terms of developing similar products.

References

- Arikunto, S. (2009). *Dasar-Dasar Evaluasi Pendidikan Edisi Revisi*. Jakarta: Bumi Aksara.
- Aritonang R., Lerbin R. (2008). Validitas dan Reliabilitas Instrumen. *Akademika: Jurnal Pendidikan Universitas Tarumanegara*, 10 (2), 159-180.
- Ariyana, Y., Pudjiastuti, A., Bestary, R., & Zamroni. (2018). *Buku Pegangan Pembelajaran Berorientasi pada Keterampilan Berpikir Tingkat Tinggi*. Kementerian Pendidikan dan Kebudayaan.
- Bagiyono. (2017). Analisis Tingkat Kesukaran dan Daya Pembeda Butir Soal Ujian Pelatihan Radiografi Tingkat 1. *Widyanuklida*, 16 (2), 1-12.
- 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>
- Candiasa, I M. (2010). *Pengujian Instrumen Penelitian Disertai Aplikasi ITEMAN dan BIGSTEPS*. Singaraja: Undiksha Press.
- Erfianti, L., Istiyono, E., & Kuswanto, H. (2019). Developing Lup Instrument Test to Measure Higher Order Thinking Skills (HOTS) Bloomian for Senior High School Students. *International Journal of Educational Research Review*, 4(3), 320–329.
- Irmawati, R. D., Supriyati, Y., & Suseno, M. (2018). Pengaruh Strategi Pembelajaran Dan Motivasi Belajar Terhadap Higher Order Thinking Skills (HOTS) Dalam Pembelajaran IPA Siswa Kelas IV Sekolah Dasar. *Tunas Bangsa Journal*, 5(2), 143–156.
- Japa, I. G. N dan Suarjana, I. M. (2015). *Buku Ajar : Pendidikan Matematika I*. Singaraja: Universitas Pendidikan Ganesha.
- Koyan, I. W. (2011). *Statistik Pendidikan*. Singaraja : Universitas Pendidikan Ganesha Press.
- Kemdikbud. (2016). Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 21 Tahun 2016 tentang Standar Isi Pendidikan Dasar dan Menengah. Jakarta: Kementerian Pendidikan dan Kebudayaan
- Kusuma, M. D., Rosidin, U., Abdurrahman, & Suyatna, A. (2017). The Development of Higher Order Thinking Skill (HOTS) Instrument Assessment In Physics Study. *IOSR Journal of Research & Method in Education*, 7(1), 26–32. <https://doi.org/10.9790/7388-0701052632>
- Nurhasanah, N., & Yarmi, G. (2018). Workshop Pengembangan Higher Order Thinking Skill (HOTS) Melalui Penerapan Pendekatan Saintifik Dalam Pembelajaran Di SDN Beji 1 Depok Jawa Barat. *Prossiding Seminar Dan Diskusi Pendidikan Dasar*, 523–528.
- Riadi, A., & Retnawati, H. (2014). Pengembangan Perangkat Pembelajaran untuk Meningkatkan HOTS pada Kompetensi Bangun Ruang Sisi Datar Developing Learning Kit to Improve HOTS for Flat Side of Space Competence. *Jurnal Pendidikan Matematika*, 9(2), 126–135. <http://journal.uny.ac.id/index.php/pythagoras>
- Rochmad. (2012). Desain Model Pengembangan Perangkat Pembelajaran Matematika. *Jurnal Kreano FMIPA UNNES*, 3 (1), 59-72.
- Sugiyono, (2011). *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.

- Surapranata, Sumarna. (2010). *Analisis, Validitas, Reliabilitas, dan Interpretasi Hasil Tes*. Bandung: Remaja Rosdakarya.
- Tanujaya, B. (2016). Development of an Instrument to Measure Higher Order Thinking Skills in Senior High School Mathematics Instruction. *Journal of Education and Practice*, 7(21), 144–148.
- Widihastuti, & Suyata. (2014). Model AFL untuk Meningkatkan Pemahaman dan Higher Order Thinking Skills Mahasiswa Vokasi Bidang Busana. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 18(2), 275–289.