Development of Higher Order Thinking Skills (HOTS) Objective Tests on Lingkungan Sahabat Kita Theme in Grade V Elementary School

Dewi Ramadhani[1], Sahyar Sahyar[2], Zulkifli Matondang[3]

{dewiramadhani47@gmail.com[1]}

Basic Education Study Program of Postgraduate School of Universitas Negeri Medan, Indonesia[1],
Physics Education Department of Postgraduate School of Universitas Negeri Medan, Indonesia[2],
Building Engineering Department, Faculty of Engineering, Universitas Negeri Medan, Indonesia [3]

Abstract. The study aims to develop Higher Order Thinking Skills (HOTS) instrument test in Lingkungan Sahabat Kita theme to standart qualifications of good test based in validity, reliability, difficulty index, discrimination index and distractor effectiveness. The resulting product is HOTS objective test instrument include cognitive dimention in Anderson and Krathwohl's taxonomy. The Type of research is development research using the ADDIE model. The data analysis technique used is qualitative and quantitative. The result of qualitative test analysis show that quality of the test was good with percentage of validity is 98.53%. The result of quantitative test analysis of 50 items obtained 38 items was valid with reliability in 0.917, the average discrimination index and effectiveness of distractor in the good category. The difficulty index in the medium category. Based on these results was concluded that the objective HOTS test developed on Lingkungan Sahabat Kita theme has good standard test qualifications and can be used as an instrument for learning outcomes on that theme.

Keywords: Higher Order Thinking Skills (HOTS), Water Cycle, Objective test, Lingkungan Sahabat Kita Theme.

1 Introduction

An educator in the learning process is not only responsible for transferring knowledge and achieving learning targets. Teachers are also required to be able to develop assessments, process evaluations and learning outcomes as guidelines for improving the learning process in the future.

The principle of assessment is used as a tool to improve the quality of the learning process and the ability of students. The assessment is carried out with an instrument that is in accordance with the abilities to be measured, both knowledge, attitudes and skills of students. Knowledge assessment is carried out to determine the level of student's ability with the material being studied [1] which includes memorization, understanding and application (Low Order Thinking Skills / LOTS) analysis, evaluation and creating (High Order Thinking Skills/HOTS) [2],[3].

HOTS is a very important skill to apply because students need to be equipped with skills in thinking and managing information to solve real-life problems. This is in line with [4]; [5] which explains that the application of HOTS in learning will cause students to get used to analyzing and being creative in solving problems found in everyday life.

Therefore, students need to be trained and faced with questions that require unusual thinking processes. But in reality, the test instrument used by the teacher is still limited to measuring the level of knowledge so that it does not develop students' thinking skills. [6]; [7] who explained that test questions in schools tend to only test students' level of knowledge (C1) and understanding (C2). The evaluation questions used did not identify students' higher-order thinking skills [8] so students were not trained to solve HOTS-based questions. [9] explained that an effective test must be able to develop higher-order thinking skills.

Based on the results of field studies, information was obtained that the application of HOTS in test questions was still considered difficult for teachers to implement. Teachers are still not fully able to develop HOTS-based test questions. So that the test questions used are still taken from books and developed by themselves which are adapted to the material that has been taught. Only 20% of the questions made by the teacher use operational verbs in the HOTS category but the questions are not classified as HOTS. In addition, there is no further analysis of the test questions related to the quality that meets the standards. [10] explains that in learning teachers tend to be less in-depth in reviewing the material and have not involved students in the process of thinking independently. So it has an impact on the ability of students to solve the problems they face.

[11] states that the ability of teachers in preparing test questions is not good. Only 30% of teachers can compile test questions and the rest take questions from the internet. [12] also state that the test questions obtained by teachers from the internet are not of good quality because to get a good and accurate test instrument takes a lot of qualitative and quantitative testing so it will take quite a long time. The preparation of the HOTS-based test instrument can be done in the form of a written test. The test is used to test students' abilities based on higher-order cognitive thinking processes developed by [13] namely analyzing (C4), evaluating (C5), and creating (C6).

Natural Sciences (IPA) is one of the basic competencies contained in thematic learning in elementary schools. Science learning aims to develop knowledge, attitude, and skill competencies through a series of logical and systematic observation processes to understand various phenomena that occur in nature. [14]; [15] explained that science learning aims to encourage students to be able to analyze material through critical, logical, and rational thinking processes. The nature of science which is a product, process, and scientific attitude facilitates students to be creative, active, and have high-level thinking skills.

The preparation of the right test instrument can affect the quality of the assessment. So according to [16] the preparation of test questions instruments must refer to the guidelines for preparing a good test which include validity, reliability, objectivity, practicality, and economy. Test items must also be developed based on basic competencies and indicators of competency achievement. [8] explained that the realization of an appropriate assessment cannot be separated from the quality of the instruments used to measure students' HOTS abilities.

The development of HOTS objective test instruments developed by researchers used the

ADDIE design. ADDIE is an educational product development concept so its application makes learning student-centered, innovative, authentic, and inspiring [17].

[18] developed the HOTS class IV questions on the theme of always saving energy which consisted of multiple choice questions and essay questions. [19] have developed a HOTS assessment in class V on theme 6 (temperature and heat material) which consists of basic competencies in science, social studies, SBdP, Civics, and Indonesian in the form of essay questions. The results obtained are the test developed to have good quality.

2 Research method

This research is development research. The development product is an instrument assessment to train students' thinking skills. Development type is using ADDIE model. 5 development steps which consist of 1) Analyze, 2) Design, 3) Development, 4) Implementation and 5) Evaluation. The steps of the development test instrument with the ADDIE model are shown in the following figure 1.
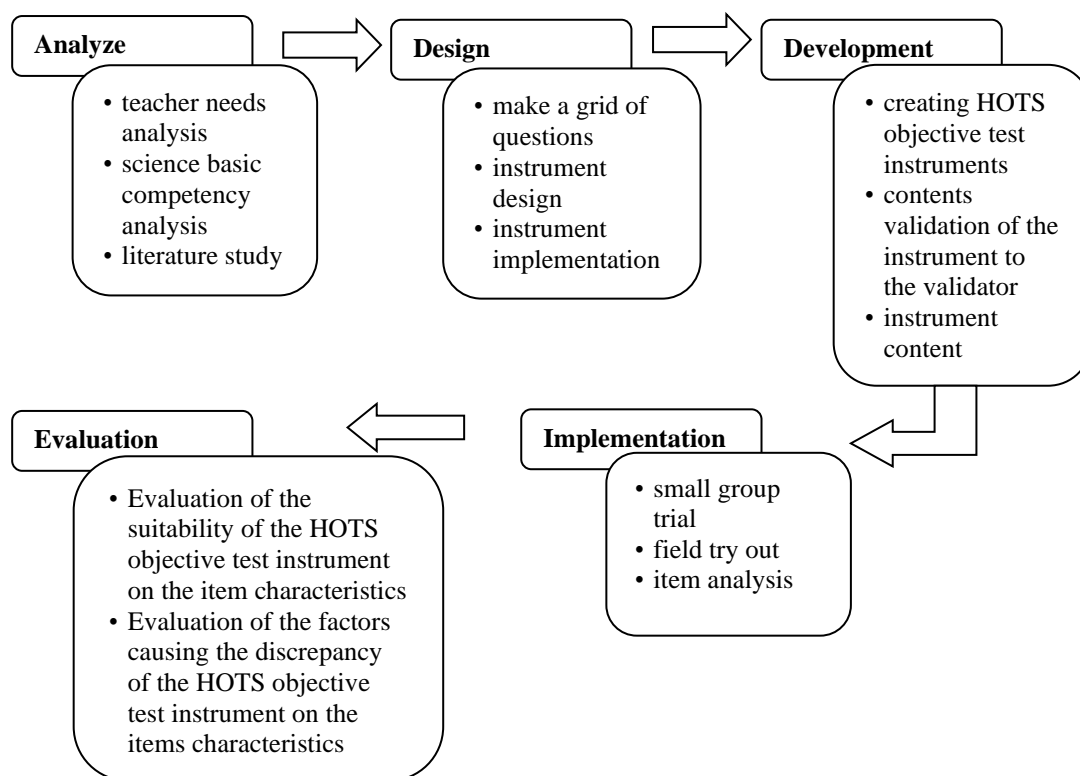

Figure 1. Steps of ADDIE Model

The instrument used in this study includes the HOTS test instrument and item validation sheet. The test item validity sheet is used to examine the test items qualitatively which contains statements regarding material, construction, and language aspects. Then, an instrument test

that has been arranged is used to do a small group trial. A small group trial was done in SDN 060877 Medan with 30 samples. Then, test results analyzed to known validity, reliability, discrimination index, and difficulty level are arranged into the initial product which is used to field tryout. Field try-out was done in MIN 3 Medan with a total of samples 94. The result of the field tryout is used to assess the instrument which has been developed as an assessment for training students in HOTS.

3 Result and discussion

A. Analyze

The results of the teacher need analysis based on interviews obtained information that the assessment of student learning outcomes was carried out with multiple choice and essay tests which were still in the cognitive realm C1 to C3. The teacher also revealed that the HOTS questions are still difficult to develop but the HOTS questions are needed to help develop students' thinking skills. The results of the basic competency analysis based on several related studies show that the development of HOTS objective questions in elementary schools is still rarely developed. Based on the results of the literature study, information was obtained that the development of HOTS questions was based on basic competencies that included operational verbs on the dimensions of the HOTS cognitive process.

B. Design

The design of the HOTS objective test instrument is to create a grid of questions that refer to indicators of achievement of learning competencies. The design of the Higher Order Thinking Skills (HOTS) test instrument which was developed in the form of multiple-choice, complex multiple-choice and true-false objective questions, consisted of 50 questions equipped with question indicators and answer keys. Based on indicators developed from basic science competencies on the theme 8 Environments of Our Friends.

C. Development

Instrument validation is done so the product developed was valid and suitable for use. The experts who asked to rate the instrument in this study were Prof. Dr. Sumarno, M.Pd., Dr. Ely Djulia, M.Pd., Dr. Hasruddin, M.Pd., Dr. Wisman Hadi, M.Hum., and senior teacher Maulidayani, M.Pd. The experts assess material, construction and language aspects. In addition, the expert also assessed the suitability of the basic competencies with the indicators on the questions. The results of the suitability of basic competencies with indicators can be seen in Table 1.

Table 1. The results of the suitability basic competencies with indicators

Indicators	Experts					Score	Validity value	Criteria
	A	B	C	D	E			
students are able to identify the benefits of water for humans, animals, and plants.	3	3	4	3	3	16	80%	Good
students are able to assess the process of the water cycle at each stage.	3	4	4	4	3	18	90%	Excellent
students are able to describe the occurrence of the water cycle.	4	2	4	4	4	18	90%	Excellent

students are able to identify the factors that affect the water cycle	3	1	4	3	2	13	65%	Good
students are able to distinguish between groundwater and surface water	1	3	3	3	3	13	65%	Good
students are able to determine the factors that affect the availability of clean water	4	4	2	4	4	18	90%	Excellent
students are able to plan simple designs related to the impact of the water cycle on life.	4	4	4	4	4	20	100%	Excellent
students are able to identify the impact of the water cycle on life.	4	4	4	4	4	20	100%	Excellent
students are able to investigate the factors that affect water quality.	4	1	4	4	4	17	85%	Excellent
students are able to analyze the effect of water quality on human life.	3	3	3	3	2	14	70%	Good
students are able to identify sources of clean water availability	4	3	3	4	4	18	90%	Excellent
students are able to criticize behavior that does not maintain the availability of clean water	4	4	4	4	2	18	90%	Excellent
students are able to develop ideas about ways to maintain the availability of clean water	3	3	3	3	3	15	75%	Good
students are able to identify the benefits of clean water	2	4	4	4	4	18	90%	Excellent
Total						218	77,8%	Good

Table 1 shows the average percentage of the HOTS instrument indicators rating is 77,8% or 'good'. The suitability of the indicator with the item is calculated by the CVR index. The results of the analysis of the suitability of the indicators with the items can be seen in table 2.

Table 2. The results of the analysis of the suitability of the indicators with the items

Category	Item number	Total
Valid	2,3,6,7,8,9,10,11,12,13,18,19,20,22,23,25,26,27,28,29, 30,32,34,35,36,37,38,39,40,41,42,43,44,45,47,48,49,50	38
Revise	1,4,5,14,15,16,17,21,24,31,33,46	12
	Total	50

Table 2 shows 38 items valid and 12 items that need to revise. Based on Table 1 and Table 2, the experts agree that the HOTS objective test questions developed by researchers are following the indicators and basic competencies to be measured. A total of 12 items were suggested by experts to be improved because there were still aspects of the material, construction, and language that were not appropriate.

D. Implementation

1) Trial Small Group. Small group trials were carried out after the questions suggested by the experts were correct. The small group trial was carried out in class V SDN 060877 Medan with 30 students. Empiric analysis of quantitative data includes validity, reliability, level of difficulty, and discrimination power. Analysis results of 50 items, it is known that the valid are 41 items (82%) and the invalid 9 items (18%). The reliability coefficient obtained is 0.923. The results of the calculation of the difficulty level of the HOTS objective test items are generally in the medium category. In detail, the calculation results in the small group trial were obtained 9 items (18%) which were in the difficult

category, 40 items (80%) were in the medium category, and 1 item (2%) was in the easy category. In detail, the discrimination power of 17 items is in the very good category, 20 items are in a good category, 8 items are in enough category, and 5 items are in the poor category. Analysis results of the small group trial can be seen in Figure 2 and Figure 3.
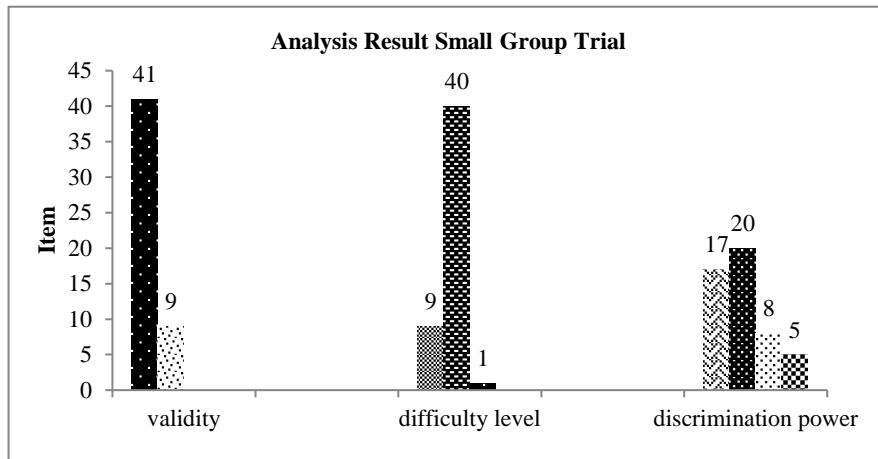

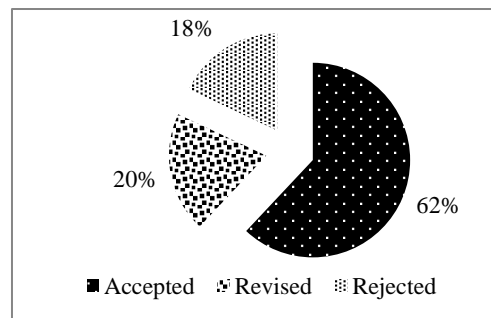**Figure 2.** Analysis result small group trial graphic

Figure 3. Empiric result small group trial

2) Large Group Trial. The large group trial was carried out in class V MIN 3 Medan with a total of 94 students using 41 questions from the results of analysis and revision in small groups. Empiric analysis of quantitative data includes validity, reliability, level of difficulty, and discrimination power. Analysis results of 41 items, it is known that the valid are 38 items (92.6%) and invalid are 3 items (7.4%). The reliability coefficient obtained is 0.917. The results of the calculation of the difficulty level of the HOTS objective test items are generally in the medium category. In detail, the calculation results in the large group trial were obtained 39 items (95.1%) were in the medium category and 2 items (4.9%) were in the easy category. In detail, the discrimination power of 21 items is in the very good category, 11 items are in a good category, 6 items are in enough category, and 3 items are in the poor category. Analysis results of the large group trial can be seen in Figure 4 and Figure 5.
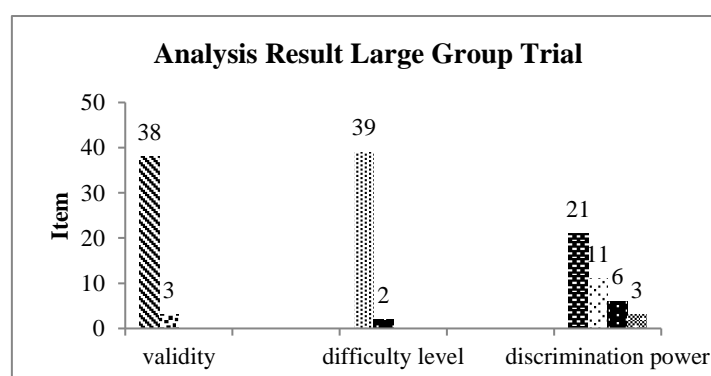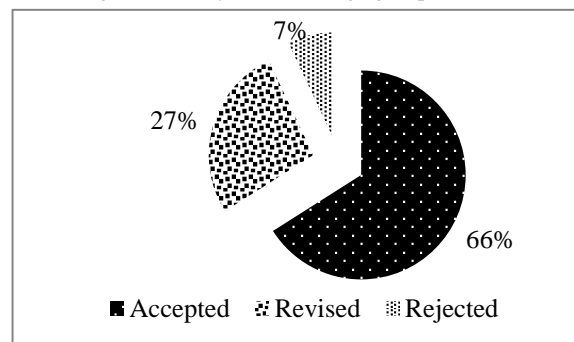
Figure 4. Analysis result large group trial

Figure 5. Empiric result large group trial

E. Evaluation

HOTS objective items that match the characteristics of the items must meet the requirements of validity, reliability, level of difficulty, and distinguishing power in the good category. The results of the small group trial obtained questions that match the characteristics of good items 41 items of 50 items tested. The results of the large group trial were obtained 38 items by the characteristics of good questions and stored in the question bank of the Higher Order Thinking Skills (HOTS) Objective Test on the Environmental Our Friends Theme in Elementary School.

Discussion

This research is the development objective test of Higher Order Thinking Skills (HOTS) on the Environmental Our Friends Theme in class V Elementary School. Development of HOTS objective instrument test through five stages namely Analyze, Design, Development, Implementation, and Evaluation.

The results of the analysis of the HOTS objective test items have a validity score of 98.53% including the very high category. The validity criteria were obtained from the results of the instrument experts' assessments on aspects of the material, construction, and language in line with research [20]; [21]. The instrument is valid if the results of the expert assessment are calculated using the Content Validity Ratio (CVR) formula that meets the specified critical value limit [22]. Based on the results of small group trials, the items developed in the valid

category were 41 items (82%) and 9 items (18%) were in the invalid category. In the large group trial, there were 38 items (92.6%) valid questions and 3 items (7.4%) invalid items. The validity of the two test groups is included in the very high category because more than 80% of the questions are valid.

Small group trials obtained a reliability of 0.930. In the large group trial, the reliability was 0.918. This result is included in the category of "very high" reliability and is in line with research [22]. According to [23] the value of the reliability coefficient ranging from the medium to a high category means that the test instrument is reliable in testing students' higher-order thinking skills even though it is used at different times the results will be the same or close to the same.

Good items are questions whose level of difficulty is in the medium category or not too easy and not too difficult [24]. Furthermore [12] explains that questions that are too easy do not stimulate students to optimize their ability to solve problems. Problems that are too difficult will make students despair and do not want to solve them because they are beyond their abilities. The results of data analysis in the small group trial obtained 9 items (18%) including the difficult category, 40 items (80%) including the medium category, and 1 item (2%) including the easy category. Difficult questions will be corrected by looking at the language aspect and the effectiveness of the distractors. The results of the large group trial data analysis obtained 39 items (95.1%) including the medium category and 2 items (4.9%) including the easy category. Based on these results, the average difficulty level of the developed objective test items is 0.48, which is in the medium category. These results are in line with research [25] which obtained an average HOTS objective test difficulty level of 0.43 and research [21] obtained an average HOTS test difficulty level of 0.33 is in the medium category.

[16] explains that items with good discriminating power mean being able to distinguish between the upper and lower groups. A good test of discriminatory power has a correlation coefficient ≥ 0.20. The discriminatory power of the items in this study was on average 0.61, this means that the distinguishing power of the HOTS objective test on the theme of Our Friendship Environment in elementary school is in a good category. This result is in line with research [21] with an average discriminating power coefficient of 0.42 and research [25] with an average discriminating power coefficient of 0.35 is in the category good.

4 Conclusion

The HOTS objective test instrument on the theme of our friend's environment in elementary school was declared valid with a validity score of 98.53%. the results of the analysis of the characteristics of the items obtained 38 (92.7%) good items and 3 items (7.3%) bad items and discarded. item has a reliability of 0.917. The difficulty level of the items obtained was 39 items (95.1%) in the medium category and 2 items (4.9%) which was in the easy category. the discrimination power of 21 items is in the very good category, 11 items are in a good category, 6 items are in enough category, and 3 items are in the poor category. Based on these results was concluded that the objective HOTS test developed on the Environmental Our Friends theme has good standard test qualifications and can be used as an instrument for learning outcomes on that theme.

References

[1] Arikunto, S. (2018). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara

[2] Wiranti, N. P.D, Suniasih, N. W., & Darsana, I. W. (2017). Pengaruh Model Pembelajaran Student Facilitator and Explaining Berbantuan Peta Konsep Terhadap Kompetensi Pengetahuan Ipa Siswa. *Journal of Education Technology*, 1(2), 204 – 210.

[3] Widiana, I. W., Parera, N. P. G., & Sukmana, A. I. W. I. Y. (2019). Media Permainan Ular Tangga Untuk Meningkatkan Hasil Belajar Siswa Kelas IV Pada Kompetensi Pengetahuan IPA. *Journal of Education Technology*, 3(4), 315–322.

[4] Sihaloho, R. R., Sahyar, S., & Ginting, E. M. (2017). The Effect of Problem Based Learning (PBL) Model toward Student's Creative Thinking and Problem Solving Ability in Senior High School. *IOSR Journal of Research & Method in Education (IOSRJRME)*, 07(04), 11–18.

[5] Sani, R. A. (2019). *Pembelajaran Berbasis HOTS (Higher Order Thinking Skill)*. Tangerang: Tsmart.

[6] Wangsa, G. N. A. S., Dantes, N., & Suastra, I. W. (2021). Pengembangan Instrumen Kemampuan Berpikir Kritis dan Hasil Belajar IPA Kelas V SD Gugus IV Kecamatan Gerokgak. *PENDASI: Jurnal Pendidikan Dasar Indonesia*, 5(1), 139–150.

[7] Puspita, N. (2018). Pengembangan LKPD HOTS Tasikgaus pada Koordinat Cartesius di Kelas VIII C SMPN 1 Bandar Dua. *Prosiding Seminar Nasional Guru Dikdas Berprestasi.* Direktorat Pembinaan Guru Pendidikan Dasar. Kementrian Pendidikan dan Kebudayaan.

[8] Hartini, P., Setiadi, H., & Ernawati. (2020). Instrumen Penilaian Berbasis LOTS dan HOTS Buatan. *Jurnal Penelitian Dan Penilaian Pendidikan (JPPP)*, *3*(1), 14–24.

[9] Jailani, Sugiman, Retnawati, H., Bukhori, Apino, E., Djidu, H., & Arifin, Z. (2018). *Desain Pembelajaran Matematika Untuk Melatihkan Higher Order Thinking Skill*. Yogyakarta: UNY Press

[10] Sihotang, L., Setiawan, D., & Saragi, D. (2017). The Effect of Learning Strategy and Self Confidence Toward Students' Learning Outcomes in Elementary School. *IOSR Journal of Reseacrh & Method in Educatiob (IOSR-JRME),* 7 (4), 65–72.

[11] Sopandi, A. (2021). Peningkatan Kemampuan Guru dalam Menyusun Soal Tes Hasil Belajar Melalui *In House Training*. Jurnal Inovasi Pembelajaran Karakter(JIPK). 6 (1), 1-9.

[12] Siregar, N., & Sahyar, S. (2021). Pengembangan Tes Objektif Pengetahuan Konseptual Fisika pada Materi Usaha dan Energi di SMA. *Jurnal Inovasi Pembelajaran Fisika (INPAFI),* 9 (1), 67–75.

[13] Anderson, L. W., Krathwohl Peter W Airasian, D. R., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *Taxonomy for Learning, Teaching and Assessing - a Revision oF Bloom's Taxonomy of Educational Objectives.* Addison Wesley Longman, Inc.

[14] Azam, I. F., & Rokhimawan, M. A. (2020). Analisis Materi IPA Kelas IV Tema Indahnya Kebersamaan Dengan HOTS. *JURNAL ILMIAH DIDAKTIKA: Media Ilmiah Pendidikan Dan Pengajaran*, *21*(1), 100–110.

[15] Dewi, S., Harahap, F., & Sutopo, A. (2021). Inquiry Based Teaching Material Development Guided in Calor ans Materials the Transfer of Class V Primary State 2 Alue Teh. *Budapest International Research and Critics in Linguistics and Education (BirLE) Journal,* 4 (1), 13–23.

[16] Sudijono, A. (2013). *Pengantar Evaluasi Pendidikan*. Jakarta: PT RajaGrfindo Persada.

[17] Branch, R. M. (2009). *Instructional Design: The ADDIE Approach*. New York: Springer.

[18] Yuliandini, N., Hamdu, G., & Respati, R. (2019). Pengembangan Soal Tes Berbasis Higher Order Thinking Skill (HOTS) Taksonomi Bloom Revisi di Sekolah Dasar. *PEDADIDAKTIKA : JURNAL ILMIAH PENDIDIKAN GURU SEKOLAH DASAR*, *6*(1), 37–46.

[19] Ulum, A.R., Hidayah, N., & Yanti, Y. (2021). Development of Assessment HOTS (Higher Order

Thinking Skills) Based on Troubleshooting for Class V SD/MI. *JMIE: Jurnal Pendidikan Madrasah Ibtidaiyah.* 5 (1), 15 - 29.

[20] Sabani, Tanjung, Y.,& Bunawan, W. (2018). Making of Test Instrument Based Higher Order Thinking Skills on Basic of Mechanic and Heat. AISTSSE 2018, DOI 10.4108/eai.18-10-2018.2287415

[21] Setiawati, W., Asmira, O., Ariyana, Y., Bestary, R., & Pudjiastuti, A. (2019). *Buku Penilaian Berorientasi Higher Order Thinking Skills*. Direktorat Jenderal Guru dan Tenaga Kependidikan, Kementerian Pendidikan dan Kebudayaan.

[22] Nursa'dah, R., Sahyar., & Bunawan, W. (2021). Development of Objective Test for High-Level Knowledge of High School Physics Materials on Temperature and Heat. *Proceedings of the 6th Annual International Seminar on Transformative Education and Educational Leadership* (AISTEEL 2021). Advances in Social Science, Education and Humanities Research, vol 591, 72 - 78.

[23] Ningsih, M.S., Bunawan, W.,& Ramadhani, I. (2020). Development of High Order Thinking Skill Test Instrument on Optical Materials. *Indonesian Physics Education Research.* 01 (1), 43 - 52.

[24] Hartati, N., & Yogi, H. P. S. (2019). Item Analysis for a Better Quality Test. *English Language in Focus (ELIF),* 2(1), 59–70.

[25] Harahap, W.Y., & Sahyar. (2019). Pengembangan Tes Objektif HOTS Materi Suhu dan Kalor di SMA/MA. *Jurnal Ikatan Alumni Fisika Universitas Negeri Medan.* 5 (3), 7 - 14.