



LOTS and HOTS Capabilities as Assessment Instruments Made by Elementary School Teachers in Jakarta

Ernawati^{1*}, Hari Setiadi¹, Puji Hartini¹

¹Program Studi Penelitian dan Evaluasi Pendidikan SPs UHAMKA, Jl. Warung Jati Kalibata, Jakarta Selatan, DKI Jakarta, Indonesia, 12740

*Correspondent Email: ernawati.pep@uhamka.ac.id



ARTICLE INFO

Article history:

Received: 12 Nov 2020

Accepted: 23 Jun 2021

Published: 30 Aug 2021

Keywords:

Assessment Instruments;
Elementary School Teacher;
HOTS;
LOTS;

ABSTRACT

Background: Assessment is a crucial step for determining whether or not a learning process has been successful. High-order thinking skills are a requirement for learning achievement in the twenty-first century. This project aims to put quantitative evaluation instruments created by Jakarta elementary school teachers to the test. **Methods:** A mixed-method is used, which combines descriptive qualitative research with document analysis and quantitative research with instrument testing on reproductive system material, which is then examined using Anates. The participants in this study were eight Jakarta-based grade VI primary school teachers. **Results:** It demonstrates that teachers' capacity to construct evaluation instruments varies widely. Testing using Anates shows 15% of the questions compiled are valid and 85% invalid, one school with strong reliability and seven other schools are less reliable. The difference shows that 42.5% are accepted, and 57.5% are rejected. The difficulty level showed that as many as 8.75% of the questions were in the very easy category, 18.75% easy, 51.25% moderate, 7.5% difficult and 13.75 very difficult. These results form the basis of the importance of the hierarchy of teacher understanding in-depth, especially in relation to the preparation of HOTS-based assessment instruments because the quality of learning success is largely determined by the assessment instrument used. **Conclusions:** The LOTS and HOTS-based capability instruments made by the teacher at 8 East Jakarta Elementary Schools have not been maximized and need development by these teachers.

Kapabilitas LOTS dan HOTS Instrumen Penilaian Buatan Guru Sekolah Dasar di Jakarta

ABSTRAK

Background: Penilaian merupakan proses penting yang menjadi tolok ukur keberhasilan proses pembelajaran. Kemampuan berpikir tingkat tinggi merupakan bagian urgensi dalam pembelajaran abad 21 yang menjadi tuntutan ketercapaian pembelajaran. Penelitian ini bertujuan untuk menguji secara kuantitatif dan mendeskripsikan instrumen penilaian yang dibuat oleh guru SD di Jakarta. **Metode:** Menggunakan mix method dengan menggabungkan penelitian kualitatif deskriptif melalui teknik pengumpulan data menggunakan analisis dokumen dan penelitian kuantitatif dengan pengujian instrumen pada materi sistem reproduksi yang kemudian dianalisis dengan menggunakan Anates. Subjek dalam penelitian ini adalah 8 guru SD kelas VI di Jakarta. **Hasil:** Menunjukkan adanya variasi kemampuan guru dalam mengembangkan instrumen penilaian. Pengujian menggunakan Anates menunjukkan 15% soal yang disusun sudah valid dan 85% tidak valid, 1 sekolah dengan reliabilitas kuat dan 7 sekolah lainnya kurang reliabel. Daya bedanya menunjukkan 42,5 % diterima dan 57,5 % ditolak. Tingkat kesukaran menunjukkan sebanyak 8,75 % soal dalam kategori sangat mudah, 18,75% mudah, 51,25 % sedang, 7,5% sukar dan 13,75 sangat sukar. Hasil tersebut menjadi dasar pentingnya hierarki pemahaman guru secara mendalam, terutama terkait dengan penyusunan instrumen assesmen berbasis HOTS karena kualitas keberhasilan pembelajaran sangat ditentukan oleh instrumen assesmen yang digunakan. **Kesimpulan:** Kapabilitas instrumen penilaian berbasis LOTS dan HOTS buatan guru pada 8 Sekolah Dasar Dikakarta Timur belum maksimal dan perlu adanya pengembangan oleh guru.

Kata kunci:

Guru SD;
HOTS;
LOTS;
Instrumen Penilaian;



© 2021 by authors. License BIOEDUSCIENCE. UHAMKA, Jakarta. This article is open access distributed under the terms and conditions of a Creative Commons Attribution (CC-BY) license.

Introduction

The standard of competence that learners must possess at the basic level is listed in the Regulation of the Minister of Education and Culture number 21 of 2016 on the standard of primary and secondary education content, while the skills that must be possessed are thinking and acting skills including creative, productive, critical, independent, collaborative, and communicative in a clear, systematic, logical and critical language, in aesthetically pleasing work, movements that reflect a healthy child, and actions that reflect the child's behaviour according to his stage of development.

The world of education continues to get the spotlight especially related to success, even effective processes in learning. In the process, an educator is highly required to be able to prepare various and innovative learning facilities so that it is expected that education can motivate learners to learn and achieve so as to improve personal, school and education qualities globally. This can be seen from the results of the effective assessment by educators of the learning process implemented.

Marzano & Pickering, (1994) in the HOTS assessment playbook (2019) explained that in the dimensions of the way students think and act students are guided to have the ability to think critically, creatively and self-regulate in thinking. These learning processes are oriented towards the quality of education. Improving the quality of learners is done by improving the quality of learning oriented to high-level thinking skills. The quality of learning also needs to be measured by assessing higher order thinking skills (HOTS). It is oriented to the demands and adaptation of 21st century learning and must be a characteristic in the learning process that starts from the learning plan, learning process and assessment based on high-level thinking skills.

Based on the results of the UN in 2019 PUSPENDIK KEMENDIKBUD explained that students are still weak in higher order thinking skills such as reasoning, analyzing, and evaluating so that teachers must be able to carry out HOTS-based assessments so that students are familiar with problems and learning oriented to higher order thinking skills in order to be encouraged by their critical thinking skills. It can be applied in everyday assessments in learning.

The development of tests that are used as a measurement of the capability of a learning process is very important because with the development of the test tool, more and more valid information can be extracted. In accordance with the opinion of Aminoro & Daryanto, (2016) explained that a test is said to be valid if the test can precisely measure what it should measure. In the opinion of Arikunto, (2013) validity is the ability of a measuring instrument to measure its measuring goals. Maolani & Cahyana, (2016) support that validity is a quality that shows the conformity between the measuring device and the goal

to be measured/what should be measured. The intended validity is the validity of the assessment used by the teacher in the learning process that has been designed to be implemented so that the learning capability can be known. Based on the opinion of these experts, it can be concluded that the validity of the assessment is the quality of the measuring instrument on the assessment to be measured in accordance with the purpose of learning.

Aminoro & Daryanto, (2016) suggest that the validity of the contents indicates an instrument condition arranged based on the content of the evaluated subject matter arranged to measure the specific purpose of the given subject matter. In accordance Yusup, (2018) about the validity of content that focuses on providing evidence on the elements in the measuring instrument and processed with rational analysis to make the assessment easier.

Bloom, (1956); Situmorang, (2018) describes that the cognitive domain or cognitive realm emphasizes intellectual aspects, such as knowledge, understanding, and thinking skills. Bloom, (1956); Situmorang, (2018) also argues that the domain of cognition is divided into six levels of hierarchy, divided into two parts, namely low-level thinking skills consisting of knowledge and understanding, while high-level thinking skills consist of application, analysis, synthesis, and evaluation. Anderson in Situmorang (2018) adds the ability to think creates as the highest level, after the ability to evaluate into the HOTS category.

The material studied at the elementary school level includes materials that are factual or based on facts found in everyday life so it is necessary for a teacher to direct the HOTS learning process, this is found by learners in everyday life can be understood and learned quickly and pleasantly in the hope that learners are able to explore and apply high-level thinking from an early age, And can be used as a provision at the next level of education. Low-level thinking proposed by Situmorang (2018) consists of the ability to know and understand which is the most basic level of thinking of the cognitive aspect or realm. Sudjana, (2010); Prasetya, (2012) posits that the cognitive realm is a realm related to intellectual learning outcomes that includes six aspects namely knowledge or memory, understanding, application, analysis, synthesis, and evaluation. The first two aspects are called low-level cognitive knowledge and understanding. Anderson & Krathwohl, (2010) explains that a low level of thinking that includes the dimensions of the process of thinking knowing (C1) and understanding (C2) that measure factual, conceptual and procedural knowledge.

Brookhart (2010) in his book explained that higher order thinking skills are divided into three categories, namely HOTS as a transfer process, HOTS as a critical thinking and problem-solving ability. Good assessment instruments must meet the appropriate criteria, while Aminoro & Daryanto, (2016) explained that a good

measuring instrument must have validity, reliability, objectivity, practicality, and economy. Wijayanto et al., (2016) that in an assessment instrument, there needs to be curricular validity based on content or content related to the material to be measured in accordance with the curriculum, syllabus, and Learning Process Plan (RPP) then the use of language in the assessment instrument will affect the level of difficulty of the problem item that is prepared so that it must pay attention to grammar in accordance with EYD. Wijayanto et al., (2016) also explained that the use of appropriate language would facilitate students in understanding the intent of the problem well so that the assessment instruments that are prepared can measure what to be measured must be logical and empirical. Based on the description of the expert opinion, it can be concluded that qualitatively, the assessment instrument must have validity, reliability, objectivity, practicality, economics and pay attention to the logical and empirical arrangement of language so that it can be said to be a qualitatively qualified instrument.

In fact, there are still many teachers who have not implemented HOTS-based assessments, so that between the learning process and HOTS-based learning outcomes is still low. Wachyudi et al., (2015) said that the government has been trying to change the assessment of cognitive, affective and psychomotor specs but has not shown maximum results. In accordance with Nurani et al., (2019) mentioned that cognitive assessment in the 2013 curriculum is the most complicated and confusing assessment, so that the assessment carried out by the teacher is only based on the understanding and knowledge of the teacher. To overcome this, the government has pursued various strategies to implement hots-based learning processes and assessments in accordance with the demands of the progress of the times. Another factor that causes low achievement of HOTS in Indonesia, namely Indonesian students who are not used to doing HOTS problems. Many teachers find it difficult to structure HOTS problems so that they use existing and previously made problems that are still in the LOTS (Lower Order Thinking Skills) category. This factor is one of the factors for children not trained in solving HOTS-based problems.

The difficulty of teachers in understanding the differences in students' abilities is also an obstacle for teachers in preparing a learning plan that contains the assessment to be done. This difficulty affects the preparation of hots. However, the difficulties experienced by this teacher can also be caused by the difficulty of teachers in understanding how to prepare HOTS-based assessment instruments used in learning.

At the international level, there are several tests used to measure the ability of learners in the form of HOTS, such as those organized by PISA (Programme for International Student Assessment) and PIRLS (Progress in International

Reading Literacy Study), the achievements obtained by Indonesian learners are not satisfactory and only reach level two of the six levels contained in PISA. This low achievement is possible because of several factors, including the learning process or even assessment used by teachers so that learners are not familiar with the form of HOTS problems. In addition, many found teachers did not test the instrument before it was used for assessment. In addition, many found teachers did not test the instrument before it was used for assessment. Mendikbud, (2019) also explained the results of PISA findings on Indonesia's literacy ability to decrease reading scores.

Based on the description above, it is important for an educator to master the preparation of assessment instruments so that the research "LOTS Capabilities and HOTS Assessment Instruments Made by Elementary School Teachers in Jakarta" is very necessary because with the right assessment instruments it will show appropriate learning results especially related to low-level thinking skills (LOTS) and high-level thinking skills (HOTS).

Methods

Scope of Research

The research method used in this study is a mixed-method by combining quantitative research to find out the validity, reliability, different power and difficulty level of problem items and qualitative research to find out cognitive levels (LOTS and HOTS) through validation by experts. The subjects of this study were eight teachers in the 6th grade of public elementary schools in Jakarta. Data collection is carried out in January - March 2020 in the even semester of the 2019/2020 school year. Material is used as material for preparing problem items by teachers, and there is material on the theme of 7th grade 6, namely material about the reproductive system.

Research Procedure

This research begins by reviewing the library as a source of research reference. Then the researchers analyzed the problems that often arise in learning carried out by elementary school teachers in Jakarta. Then the researchers made observations to dig up information about assessments conducted by elementary school teachers in Jakarta.

The observation results are used to continue the preparation of research instruments in the form of interview guidelines related to the teacher's understanding of LOTS and HOTS-based assessment instruments and assessment applications used in learning. In-depth interviews were also conducted to support document data using assessment instruments made by elementary school teachers in Jakarta. Assessment instruments made by teachers are then validated by two experts who are competent in the assessment of IPA subjects at the

elementary level. The problem points compiled by the teacher are also tested on learners to find out the validity, reliability, different power, and difficulty level of each problem item. The material that is used as material for the preparation of problem items is the material on theme seven about the reproductive system.

Data Collection and Data Analysis

Data collection is done by collecting documents in the form of assessment instruments made by teachers and testing problem points on learners. Quantitative data analysis is done using the Anates application, and qualitative data analysis is analyzed with comparative analysis, namely by comparing expert validation results, bloom taxonomic theory concepts and interview results with 6th-grade elementary school teachers.

Results

The first research results in the form of documents in the form of assessment instruments made by teachers are validated by two experts, the results of which can be seen in the following table.

Table 1. Expert Validation 1

School	LOTS	HOTS	Other
School A	6	4	0
School B	8	2	0
School C	8	2	0
School D	7	3	0
School E	10	0	0
School F	8	2	0
School G	5	5	0
School H	10	0	0
Sum	62	18	0

Two experts also validate the assessment instrument compiled by the teacher. The results of expert validation two can be seen in Table 2 below:

Table 2. Expert Validation 2

School	LOTS	HOTS	Other
School A	9	1	0
School B	6	4	0
School C	9	1	0
School D	6	4	0
School E	4	4	2
School F	8	1	1
School G	7	3	0
School H	10	0	0
Sum	59	18	3

Data on instrument test results provided to learners is quantitatively analyzed with anas software whose results can be seen in the following table:

Table 3. Results of Analysis with Anates Software

School	Validity		Reliability	Difficulty level		Different power	
	V	IV		TM	TL		
A	2	8	0,38	6	4	8,75	SM
B	1	9	0,36	4	6	18,75	M
C	0	10	-0,04	0	10	5125	SD
D	4	6	0,75	9	1	7,5	S
E	0	10	0	2	8	13,75	SS
F	0	10	-1,18	3	7		
G	3	7	0,56	4	6		
H	2	8	-0,06	6	4		
Sum	12	66		34	46		
%	18,18	100		42,5	57,5		

Discussions

Overall in school A there are 6 points of questions that fall into the category LOTS and 4 points of questions included in the HOTS category, There is a difference in rationalization between constituents and experts is possible because of different levels of understanding, Based on interviews with teacher A shows that LOTS-based assessment instruments are easy problems while HOTS-based assessment instruments are difficult problems, so that the constituent is not right in using KKO for indicators about it.

As supporting data from the results of the analysis of the problem compiled by teacher A in school A was tested using Anates software while the result was only two valid questions, namely at numbers 4 and 10 while eight other problems were invalid, the quality of instruments in school A showed weak reliability with a reliability of 0.38. While the other power there are 6 points of questions accepted and 4 points of the problem rejected, to the difficulty there are 2 points of very difficult questions, 2 points of simple questions, 1 point of very simple questions and 5 points of moderate problems. Proportionally from the level of difficulty of the instrument made by teacher A shows the right proportion because each category has a problem item that represents the level of difficulty. Proportionally in school, A is proportional because there are already HOTS problem points in the assessment device. But there needs to be an improvement in content so that more quality instruments such as research conducted by Rudhito and Prasetyo need trials and revise the practical aspects of the problem developed for assessment.

In school B, the results of the analysis showed that there were 8 points of questions that fall into the LOTS category and 2 points of questions fall into the HOTS category, which is then coupled with the data of the results of the analysis with Anates, which shows that the valid problem is only one point of the problem and nine other questions are invalid with the reliability of 0.36 with weak categories. But in terms of content, the problem item created is in accordance with the validity of the content because the

material to be measured in the problem item is in accordance with the achievements on the indicator. Then there are 4 points of questions received and 6 points of questions rejected based on the different power. Furthermore, the level of difficulty in the instrument arranged there are 2 points of very difficult questions and 6 points of moderate problems and 2 points of simple problems. The number of more dominant invalid problem items is something that needs to be considered by teachers in school B. In contrast to the obstacles that arise in the research conducted by Ratnawati that teachers have difficulty understanding the speaker's material delivered during training because of differences in perception from the speaker to cause confusion when applying HOTS-based assessments.

In school C, the number of questions based on the cognitive domain includes two questions in the HOTS category, and eight other problem items fall into the LOTS category. Then the results of the analysis with Anates 10 points of invalid problems with reliability is very weak even worth -0.04 , which indicates that the problem must be revised in total even though the content has fulfilled the full validity of the content. Then the difficulty level of the problem there is 2 very difficult questions, 3 points of medium questions, three simple questions and 1 point of very simple questions. They are judging from the different power all about being rejected. This means that the instruments that are compiled must be reviewed and repaired so that they can be used as good assessment instruments. The ability of learners is also the cause of the lack of functioning of the problem points compiled. This is similar to the results of research conducted by Wulandari et al which states that different student learning achievement causes teachers to make different assessments with the planning that has been made.

In school D, expert validation results showed that 7 points of questions belonged to the LOTS category and 3 points of the problem belonged to the HOTS category. In school D, the results of the analysis using Anates showed that the 10 points of questions compiled had a strong reliability level with a number of 0.75 with the number of 4 questions already valid, namely number 2,6,8,9 and 6 other problems were invalid. Based on the difference, there are 9 points of questions accepted, and 1 point of the question rejected, while based on the difficulty level 1 point of a difficult question, 7 points of the medium problem and 2 points of a simple problem. The assessment process needs to be not continued to the maximum so that the results will be in accordance with what is expected as described by Alimuddin that the assessment can be used as a mapping of learners' learning difficulties and improvement of the learning process.

Analysis of problem items in school E as a whole the details of the problem that have been analyzed and

validated by experts turned out that all the points of the problem compiled by teacher E fall into the LOTS category. Supporting data used from Anates analysis shows that 10 problems compiled are not valid with a reliability of 0 and can be said to be very weak, with 8 points of questions rejected and 2 points of questions accepted based on different forces. Then based on the level of difficulty, there are 2 points of very difficult questions, 1 point of difficult questions, 5 points of moderate problems, 1 point of simple questions and 1 point of very simple questions. The content of teachers don't pay attention to the suitability of the problem points with the indicators compiled, so it needs a self-evaluation from the teacher. This is in accordance with Hadiana's statement stating that there needs to be an internal assessment that must be done by the classroom teacher as material for ongoing self-evaluation that can be done by stating that hots-based problems must be seen from several issues that are not valid can be used as improvements on various instrument building factors, this is in accordance with Hartini and Sukarjo who state that HOTS-based problems must be seen from several issues. What are aspects such as material, construction and language.

Ten points of questions compiled in school G turned out to be 7 points of questions that fall into the lots category and 3 points of questions included in the HOTS category. Teachers at G school have also participated in preparation training on HOTS. Then the results of the analysis showed 2 points of the same problem according to the teacher and both experts, namely items number 2 and 4 or 20% of this difference is possible because the understanding gained by the teacher during training is not applied optimally in daily assessments. In school G, there is 1 point of inappropriate questions about the reproductive system, namely at number 7 based on the validation of the second expert. What is conveyed by the teacher during the interview is that it is necessary to procure further training to establish self-understanding related to HOTS-based assessments. This obstacle can be answered from the exposure delivered by Gusmarni that workshop activities can improve the competence of teachers in compiling problem points in elementary school.

Analysis of the subject points in school H showed that all the problems analyzed by experts fall into the LOTS category with a very weak reliability of -0.06 based on analysis with Anates. Then the validity yes there are 2 valid questions and eight invalid questions, 6 points of questions accepted and 4 points of questions rejected based on the different power. Next is the difficulty level, 1 point of a difficult question, 4 points of the medium question, 3 points of a simple question.

Overall, the problem items created by teachers who fall into the LOTS category amount to 81.25% of the total and 18.75% of the problem items are included in the HOTS

category. This is similar to the research conducted by Samosir et al., (2019) with the results of the number of HOTS quality questions amounting to 51% and LOTS amounting to 49%, the difference in the number of HOTS in this study is less than the research of Aldenan et al. In contrast to the research conducted by Himmah, (2019) where in his research also analyzed about mot's levels with the questions analyzed in his research is about PAS in MTK subjects. But the method used in his research is the same as this study, namely with descriptive analysis. Other similar research on the analysis of the point of the question was also conducted by Cahyono & Adilah, (2016) and Muklis & Oktora (2015), who used the categories of cognitive level knowing, applying and reasoning.

Data analysis using Anates software to determine the validity, reliability, different power and difficulty. The results of 30 questions there are 14 questions already valid 16 other problems are invalid. After an interview with elementary school teachers in East Jakarta, the problem of educators neglecting their duties and functions such as not doing an analysis of the problems given to students.

Proportionally the hot's problem is arranged with fewer numbers because each assessment device must be adjusted to the work time so that there are no significant obstacles for learners who work. In addition, the teacher still has difficulty in compiling hot's problems. It is possible that the teacher has not participated in the preparation of HOTS.

Based on interviews with 6th-grade teachers, all pay attention to the validity of the content when compiling problem items because the suitability of the material becomes the main benchmark. Then associated with the results of validation by experts, the validity of the content contained in each point of the problem has also been but what needs to be improved on the assessment instrument is KKO (Operational Verb) on indicators that do not match the form of questions that arise in the problem.

Related to the quality of the points about all teachers who were the subject of the study said that never tested the validity of reliability, difficulty level, different power on the grounds of class administration demands that took longer, so in this study tested on every instrument made by the teacher to find out the quality of the teacher's problem. Testing is done with the help of Anates software to find out the validity, reliability, different power, difficulty level. As for the results that have been listed in the recapitulation table of the analysis results. This study obtained the interesting fact that the questions provided by the government in the books published by the ministry, banks of problems that are easily accessible online and bank problems in printed books turned out to be the main source of teachers in testing their students in daily repeats, midterms, end-of-semester exams and even end-of-year assessments, so that the ability of teachers who have been owned from training is less developed in daily life.

Furthermore, the problems that teachers usually use are the problems that are already available in the package book that has been provided to students. Especially daily repeats. But there are also points of questions that teachers modify in accordance with the conditions and materials delivered in the classroom. Testing of assessment instruments before being used as a measuring tool for the success of the learning process must indeed be done so that learning goals can be achieved properly, testing on this study is done as research support data to find out the quality of the problem items compiled, which in previous research conducted by Hartuti & Handayani, (2019) by not testing so that they only know the results of the analysis only. In his research, it was produced that the implementation of the 2013 curriculum assessment, in general, was in accordance with the 2013 curriculum assessment standards with the creation of HOTS problems from UH PTS, PAS was in accordance with syllabus, RPP, teacher's book, and curriculum standards 2013.

Conclusion

Based on the results of the discussion above, it can be concluded that the instruments made by elementary school teachers exist that have met the requirements of LOTS and HOTS-based assessments proportionally. Still, there are assessment instruments that only measure lots, and there are errors in determining the cognitive level used in compiling problem points in their assessment devices. The results of the analysis data using Anates showed low validity, reliability, different power and difficulty levels of problem items compiled by elementary teachers.

Based on this study, the advice that can be given is that teachers should understand the difference in cognitive levels (C1-C6) so that they will be more regular in compiling lots and HOTS-based assessment instruments that are proportional. The applicable implication is the preparation and use of LOTS and HOTS-based assessment instruments.

Declaration statement

The authors reported no potential conflict of interest.

References

- Aminoro, & Daryanto. (2016). *Evaluasi dan Penilaian Pembelajaran Kurikulum 2013*. Gava Media.
- Anderson, & Krathwohl. (2010). *Kerangka Landasan untuk Pembelajaran. Pengajaran dan Assesmen*. Pustaka Pelajar.
- Arikunto, S. (2013). *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara.
- Bloom, B. S. (1956). *Taxonomy of Educational Objectives : The Classification of Educational Goals, Handbook I Cognitive Domain*. Longmans, Green and Co.
- Brookhart, S. M. (2010). *How to Asses Higher-Order*

Thinking Skills in Your Classroom. ASCD.

- Cahyono, B., & Adilah, N. (2016). Analisis Soal dalam Buku Siswa Matematika Kurikulum 2013 Kelas VIII Semester I Berdasarkan Dimensi Kognitif dari TIMSS. *JRPM (Jurnal Review Pembelajaran Matematika)*, 1(1), 86–98. <https://doi.org/10.15642/JRPM.2016.1.1.86-98>
- Hartuti, M., & Handayani, D. E. (2019). Analisis Penilaian Kognitif Kurikulum 2013 Kelas Rendah MI Sabilul Ulum Mayong Jepara. *El-Ibtidaiy: Journal of Primary Education*, 2(1), 1–8. <https://doi.org/10.24014/EJPE.V2I1.7370>
- Himmah, W. I. (2019). Analisis Soal Penilaian Akhir Semester Mata Pelajaran Matematika Berdasarkan Level Berpikir. , 3(1). *Journal of Medives*, 3(1), 55–63. <https://doi.org/10.31331/medivesveteran.v3i1.698>
- Maolani, R., & Cahyana, U. (2016). *Metodologi Penelitian Pendidikan*. Rajawali Pers.
- Marzano, & Pickering. (1994). *Assesing Student Outcome Performance Assessment using the Dimentions of Learning Model*. Assosiation for Supervision and Curriculum Development.
- Mendikbud. (2019). *Hasil PISA Indonesia 2018: Akses Makin Meluas, Saatnya Tingkatkan Kualitas*. Kemendikbud.
- Muklis, Y. M., & Oktora, S. R. (2015). Analisis Deskriptif Soal-Soal dalam Buku Siswa Kurikulum 2013 (Edisi Revisi) dan BSE Pelajaran Matematika SMP Kelas VII Ditinjau dari Domain Kognitif TIMSS 2011. *Prosiding Seminar Nasional Pendidikan Matematika UMS 2015*.
- Nurani, H., Artharina, F. P., & Kiswoyo. (2019). Analisis Pelaksanaan Penilaian Kognitif Berbasis Kurikulum 2013 Sabiul Ulum Mayonglor Kabupaten Jepara. *Indonesian Journal Of Educational Research and Review*, 2(2), 172–181. <https://doi.org/10.23887/ijerr.v2i2.17625>
- Prasetya, T. I. (2012). Meningkatkan Keterampilan Menyusun Instrumen Hasil Belajar Berbasis Modul Interaktif Bagi Guru-Guru Ipa Smp N Kota Magelang. *Journal of Educational Research and Evaluation*, 1(2), 106–112.
- Samosir, A., Hasruddin, H., & Dongoran, H. (2019). Analisis Kuantitas dan Kualitas Pertanyaan Guru Biologi dan Siswa Materi Sistem Ekskresi. *Jurnal Pelita Pendidikan*, 7(1), 009–015. <https://doi.org/10.24114/JPP.V7I1.10523>
- Situmorang, J. (2018). *Higher Order Thinking Skills*. MDP Media.
- Sudjana, N. (2010). *Penilaian Hasil Proses Belajar Mengajar*. Remaja Rosdakarya.
- Wachyudi, I., Sukestiyarno, & Waluya, B. (2015). Pengembangan Instrumen Penilaian Unjuk Kerja Pada Pembelajaran Dengan Model Problem Solving Berbasis Tik. *Journal of Research and Educational*

Research Evaluation, 4(1), 20–27.

Wijayanto, P. A., Allifah, A., & Amirrudin, A. (2016). Evaluasi Kualitas Instrumen Tes dalam Pembelajaran Geografidi MAN 2 Kota Batu . *Jurnal Geografi : Media Informasi Pengembangan Dan Profesi Kegeografian*, 13(2), 102–224.

Yusup, F. (2018). Uji Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif. *Jurnal Tarbiyah : Jurnal Ilmiah Kependidikan*, 7(1), 17–23. <https://doi.org/10.18592/tarbiyah.v7i1.2100>