

Cognitive domain analysis (LOTS and HOTS) assessment instruments made by primary school teachers

Puji Hartini; Hari Setiadi; Ernawati*

Universitas Muhammadiyah Prof. Dr. HAMKA

Jl. Limau II, Kramat Pela, Kebayoran Baru, Kota Jakarta Selatan, Jakarta 12130, Indonesia.

*Corresponding Author. E-mail: ernawati.pep@uhamka.ac.id

ARTICLE INFO

Article History

Submitted:

12 September 2020

Revised:

23 January 2021

Accepted:

25 January 2021

Keywords

cognitive domain analysis;
assessment instruments;
science

Scan Me:



ABSTRACT

This research aims to qualitatively analyze the validity of items and the suitability of cognitive domains (LOTS and HOTS) assessment instruments on natural science subjects made by elementary school teachers in East Jakarta. The method used is descriptive qualitative method with analysis of observations, documents in the form of teacher-made assessment instruments, interviews, and results of expert validation which are analyzed by comparison analysis techniques. The observation results show that all schools use the questions that are available in textbooks owned by students for assessment and the results of analysis of teacher-made assessment instruments validated by experts, there are 81.25% items included in the LOTS category, while 18.75% are included in the HOTS category, so it can be concluded that: (1) the instruments used by elementary school teachers in East Jakarta have fulfilled the content validity, (2) the cognitive domain (LOTS & HOTS) on the instruments used by elementary school teachers are proportional, (3) the quality of assessments conducted by elementary school teachers in East Jakarta is good with a record of improvement, (4) the implementation of assessments conducted by elementary school teachers in East Jakarta has followed the assessment standards provided by the government.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



How to cite:

Hartini, P., Setiadi, H., & Ernawati, E. (2021). Cognitive domain analysis (LOTS and HOTS) assessment instruments made by primary school teachers. *Jurnal Penelitian dan Evaluasi Pendidikan*, 25(1), 16-24.

doi:<https://doi.org/10.21831/pep.v25i1.34411>

INTRODUCTION

The competency standards that students at the basic level must have are listed in the [Regulation of the Minister of Education and Culture No. 21 of 2016](#) concerning the content standards for primary and secondary education. The skills that must be possessed are the thinking and acting skills, including creative, productive, critical, independent, collaborative, and communicative in clear, systematic, logical and critical language, in aesthetic works, movements that reflect healthy children, and actions that reflect children's behavior according to their developmental stage.

[Regulation of the Minister of Education and Culture No. 20 of 2017](#) concerning the team of performance assessor in the Ministry of Education and Culture and [Regulation of the Minister of Education and Culture No. 23 of 2016](#) concerning education assessment standards have mandated that learning outcomes assessment by educators aims to monitor the process to improve the effectiveness of learning. Entering the 21st century is marked by the rapid advancement of technology in various fields of everyday life which triggers a demand for an era that continues to move from various aspects of life, one of them is in terms of education.

The world of education continues to receive the spotlight, especially regarding success, even the effective process of learning. In the process, an educator is highly required to be able

to prepare varied and innovative learning facilities so that it is hoped that education will be able to motivate students to learn and achieve so that they can improve their personal, school and education qualities globally. This can be seen from the results of the effective assessment by educators of the learning process being carried out.

The learning process will run well, if all readiness is done well by the teacher, including the way of assessment to see the level of success and student achievement. The ability to think at a higher level will be truly measurable, if you use the right measuring tool or instrument so that it needs to be considered in the preparation of the instrument, with the aim of maximizing the expected achievement. The change in the education climate is expected to be able to produce future development candidates who are competent, independent, critical, intelligent, creative and ready to face various kinds of challenges (Mulyasa, 2017).

According to Mulyasa (2017), in the implementation of the 2013 curriculum the learning process is required to apply the HOTS-based learning concept with various innovative, creative, collaborative, problem-based and problem solving learning models. This means that the learning that has been implemented has made it possible to apply HOTS-based assessments.

Assessment, according to the Regulation of the Minister of Education and Culture No. 23 of 2016, is the process of collecting and processing information to measure the achievement of student learning outcomes. The results of this measurement process will be used as a reference for the success rate of learning. Meanwhile, education assessment standards in the scope of education assessment in primary and secondary education consist of an assessment of learning outcomes by educators; assessment of learning outcomes by educational units; and assessment of learning outcomes by the Government.

Marzano and Pickering (1997) in Setiawati et al. (2019) explain that in the dimensions of how to think and act, students are directed to have the ability to think critically, creatively and self-regulate in thinking. These learning processes are oriented towards the quality of education. One of the ways to improve the quality of students is through improving the quality of learning that is oriented towards higher order thinking skills. The quality of learning also needs to be measured by an assessment that is oriented towards higher order thinking skills (HOTS).

Based on the results of the 2019 National Examination, the Ministry of Education and Culture's center for educational assessment explains that students are still weak in higher order thinking skills, such as reasoning, analyzing, and evaluating so it is necessary for teachers to be able to carry out HOTS-based assessments so that students are familiar with the questions and learning oriented to higher order thinking skills in order to encourage critical thinking skills. This can be applied in daily assessments in learning.

The realization of an appropriate assessment is inseparable from the quality of the form of the instrument used to measure students' higher order thinking skills. In connection with the demands of the era regarding the ability to think at a high level as a step to prepare a golden generation in 2045, an educator is expected to be able to become the main force spearhead to achieve these national goals.

Regulation of the Minister of Education and Culture No. 104 of 2014 describes an assessment instrument, which is a measuring tool used to assess the learning outcomes of ideal participants with a test and attitude scale. Amirono and Daryanto (2016) explain that assessment is the application of various methods and the use of various assessment tools to obtain information about the extent to which students' learning outcomes or the achievement of students' competencies (series of abilities). It is also explained by Angelo (1991) in Amirono and Daryanto (2016) that assessment is a simple method that can use faculty (schools) to collect feedback, early and after, on how well their students learn what they teach. Then, Kurniasih and Berlin (2016) explain that assessment is a step taken for, as, and for learning. Boyer and Ewel in Amirono and Daryanto (2016) define assessment as a process that provides information about individual learners about curriculum or programs, about institutions, or everything related to the institutional system.

Based on the aforementioned experts' opinion, it can be concluded that the instrument is a measuring tool for assessing, while the assessment is the process of interpreting the measurement result data which is used to determine the level of achievement of the learning process. Thus, the instrument is a tool used to measure a measuring object with the process of data collection and data interpretation.

Amirono and Daryanto (2016) explain that a test is considered valid if it can accurately measure what it should measure. Arikunto (2011) states that validity is the ability of a measuring instrument to measure its measuring target. It is supported by Maolani and Cahyana (2015) that:

Validity is a quality that shows the suitability of the measuring instrument with the objectives to be measured/what should be measured. The validity that is meant is the validity of the assessment used by the teacher in the learning process that has been designed to be implemented so that learning achievement can be known.

Based on the aforementioned experts' opinions, it can be concluded that the validity of the assessment is the quality of the measuring instrument for the assessment to be measured in accordance with the learning objectives.

Amirono and Daryanto (2016) suggest that content validity indicates a condition of an instrument arranged based on the content of the subject matter being evaluated which is structured to measure the specific objectives of the given subject matter. It is in line with the opinion of Yusup (2018) regarding the validity of content which focuses on providing evidence on the elements that exist in measuring instruments and is processed with rational analysis so that the assessment will be easier to do. Suryanto and Sutinah (2011) restate that content validity is needed to answer the question to what extent the items in the test can measure the overall material that has been taught. Some examples of the elements that are assessed in content validity in Yusup (2018) are as follows: (1) operational definition of a variable, (2) representation of questions according to the variables to be studied, (3) number of questions, (4) answer format, (5) scale on the instrument, (6) scoring, (7) instructions for charging the instrument, (8) processing time, (9) population sample, (10) grammar, (11) writing layout (writing format).

The Minister of Education and Culture (2013) in Wardhani and Putra (2016) explains that "educational assessment must have basic principles, namely valid, objective, fair, integrated, comprehensive, sustainable, systematic based on criteria, economical, accountable and educational." Bloom (1956) in Situmorang (2018) asserts that the cognitive domain or cognitive domain are behaviors that emphasize intellectual aspects, such as knowledge, understanding, and thinking skills. Bloom et al. (1956) state that the cognitive domain is related to knowledge which involves the process of recalling specific and universal things, recalling methods and processes or recalling patterns, structures or settings. Bloom in Situmorang (2018) also argues that the cognition domain is divided into six hierarchical levels, which are then divided into two parts, namely Lower Order Thinking Skills (LOTS) consisting of knowledge and understanding. The second is Higher Order Thinking Skills (HOTS) which consists of application, analysis, synthesis, and evaluation. Anderson in Situmorang (2018) adds that the ability to think creates as the highest level, after the ability to evaluate is included in the HOTS category.

The material studied at the elementary school level includes material that is factual or based on facts found in everyday life so it is necessary for a teacher to direct the HOTS learning process, this is what students encounter in everyday life can be understood and learn with ease and fun in the hope that students are able to explore and apply high-level thinking skills from an early age, and can be used as provisions for 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 from the cognitive aspect or domain. Sudjana (2010) in Prasetya (2012) suggests that the cognitive domain is a domain related to intellectual learning outcomes which includes six aspects, namely knowledge or memory, understanding, applica-

tion, analysis, synthesis, and evaluation. The first two aspects are called low-level cognitive, namely knowledge and understanding. Besides, Anderson and Krathwohl in Pi'i (2016) explain that LOTS is a low-level thinking which includes the dimensions of knowing (C1) and understanding (C2) thinking processes that measure factual, conceptual and procedural knowledge. In addition, Brookhart (2010) in her book explains that higher-order thinking skills are divided into three categories: HOTS as a transfer process, HOTS as critical thinking skills, and HOTS as a problem solving.

Definitions that I find helpful fall into three categories: (1) those that define higher-order thinking in terms of transfer, (2) those that define it in terms of critical thinking, and (3) those that define it in terms of problem solving, terms of problem solving (Brookhart, 2010).

A good assessment instrument must meet the appropriate criteria. Therefore, Arikunto (1992) in Amirano and Daryanto (2016) explains that a good measuring instrument must have validity, reliability, objectivity, practicality, and economics. Purwanto (2011) in Wijayanto et al. (2016) believes that in an assessment instrument, there is a need for curricular validity based on content or content related to the material to be measured in accordance with the curriculum, syllabus, and Learning Process Plan, then the use of language in the assessment instrument will affect the level of difficulty of the items arranged so that it must pay attention to grammar in accordance with the correct spelling. Wijayanto et al. (2016) also explain that the use of appropriate language will make it easier for students to understand the meaning of the questions well so that the assessment instruments that are arranged can measure what they want to measure. Sudijono (1991) in Khaerudin (2015) suggests that a valid instrument must be logical and empirical. Based on the description of the expert's 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 linguistic arrangement so that it can be said to be a qualitatively quality instrument.

In fact, there are still many teachers who have not implemented HOTS-based assessments so that the learning process and HOTS-based learning outcomes are still low. Wachyudi et al. (2015) state that the government has made efforts to change the assessment on cognitive, affective and psychomotor aspects, but has not shown maximum results. It is in accordance with Nurani et al. (2019) who state that the cognitive assessment in the 2013 curriculum is the most complicated and confusing assessment so that the assessment made by the teacher is only based on the teacher's understanding and knowledge. Setiadi (2016) states that the assessment in the 2013 curriculum is considered more complicated than the assessment in the previous curriculum. To overcome this, the government has pursued various strategies to implement HOTS-based learning processes and assessments in accordance with the demands of the times. Another factor causing the low achievement of HOTS in Indonesia is that the students are not used to working on HOTS questions. Many teachers find it difficult to compile HOTS questions so they use existing and previously-made questions that are still in the LOTS (Lower Order Thinking Skills) category. This factor is one of the factors for untrained children in solving HOTS-based questions. However, in an assessment tool the number of questions presented is the comparison of the proportions between LOTS and HOTS as agreed by each school, 80% LOTS-based and 20% HOTS-based as the implementation of the assessment set in the 2013 curriculum that students must be trained and accustomed to doing questions based on HOTS. These factors also affect the achievement of HOTS ability because the one who provides an assessment to see the achievement of HOTS ability is the teacher so that when HOTS ability is low, the teacher needs to re-examine the learning process until the assessment process used includes the type of instrument used. Situmorang (2018) explains that high-order thinking skills are a competitive advantage for students, thus higher-order thinking skills need to be developed in learning, so it is important for teachers to understand correctly how to assess these abilities.

The difficulty of teachers in understanding the differences in student abilities is also an obstacle for teachers in preparing lesson plans that contain assessments to be carried out. This difficulty affected the preparation of HOTS questions. However, the difficulties experienced by this teacher can also be caused by the difficulty of the teacher in understanding how to prepare HOTS-based assessment instruments used in learning.

At the international level, there are several tests used to measure students' ability in the HOTS form, such as those held by PISA (Program for International Student Assessment) and PIRLS (Progress in International Reading Literacy Study). The achievements of Indonesian students are not satisfactory and only reached level two of the six levels contained in PISA. This low achievement is possible due to several factors, including the learning process or even the assessment used by the teacher, so students are not familiar with the HOTS questions. Assessment is part of evaluating the achievement of students and teachers in teaching (Nugroho, 2018). Based on the aforementioned description, it is important for an educator to master the preparation of assessment instruments, so that this research is very necessary to be conducted.

RESEARCH METHOD

The method used in this research is descriptive qualitative method by qualitatively describing the data obtained from the field. The research was conducted at eight state elementary schools in East Jakarta from January to February 2020. The data collection was carried out by observation, documents, and interviews, analyzed using comparative analysis techniques for checking the validity of the data. Examination of the research data must be carried out to ensure and confirm the results of the research before making conclusions. According to Guba (1981), there are four aspects of the validity or quality of qualitative research. The four aspects can be seen in Table 1.

Table 1. Aspects of Validity of Qualitative Research Perspective by Guba (1981)

Aspect	Scientific Term	Naturalistic Term
Truth Value	Internal Validity	Credibility
Applicability	Eksternal Validity Generalizability	Transferability
Consistency	Reliability	Dependability
Neutrality	Objectivity	Confirmability

Source: Guba (1981)

FINDINGS AND DISCUSSION

The findings based on observations were made at eight public elementary schools in East Jakarta which were used as research sites by entering the classroom when the teacher was teaching and giving an assessment that eight teachers at the schools only used questions available in textbooks owned by students because they were considered more practical and easy for teachers to do.

The next finding was that the results of interviews with grade XI teaching teachers turned out that there was one teacher who had not attended HOTS-based instrument preparation training, while seven teachers had attended the training but they said they had problems in its implementation because it was still something new for elementary school level teachers. This is the reason teachers do not make their own assessment instruments.

Documents taken as data were in the form of assessment instruments made by eight public elementary school teachers in East Jakarta, each of which consisted of ten multiple choice questions in science subject of grade XI with the reproductive system material so that all the items were 80 items, then are validated by two expert. The validation results are presented in Table 2. Then, the document was also validated by a second expert which can be seen in Table 3.

Table 2. The Result of Teacher Evaluation Instrument Validation by Expert 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
Total	62	18	0

Table 3. The Result of Teacher Assessment Instrument Validation by Expert 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
Total	59	18	3

Based on Table 2, there are 62 items included in the LOTS category and 18 items in the HOTS category. This categorization is based on the cognitive domain. Validation is carried out based on the suitability of the content or material to be measured, namely related to the reproductive system.

Validation conducted by the second expert in Table 3 shows that there are 59 items in the LOTS category and 18 items in the HOTS-based category, with three items that are unclear and not included in the reproductive system material. These three items were found in school E totaling two items and school F totaling one item.

Table 2 and Table 3 present different validation results, because there are different understandings between the two experts, but as a whole, they do not show a significant difference. Overall, the items made by the teacher who were included in the LOTS category were 81.25% of the total, and 18.75% of the items were included in the HOTS category. This is similar to the research that was conducted by Samosir et al. (2019) with the results of the number of HOTS quality questions of 51% and LOTS of 49%, the difference is that the number of HOTS in this study was less than that of Samosir et al. (2019) material, and the research was also carried out in different places. It is contradictory to the context of the research that was carried out by Himmah (2019), in which she also analyzed the level of the MOTS, with the end-semester assessment questions on mathematics subject. However, the method used in her research was the same as this research, namely, descriptive analysis. Moreover, other similar researches on item analysis were also conducted by Cahyono and Adilah (2016), and also Muklis and Oktora (2015) who used the cognitive level category of knowing, applying, and reasoning.

Data analysis that was conducted using Anates software to find out the validity, reliability, differentiation power, and level of difficulty in this study distinguishes it from the research that was conducted by Sudrajat (2018) which deals manually with the results of 30 questions: there are 14 valid questions, and 16 other questions are not valid. After conducting interviews with elementary school teachers in East Jakarta, it turned out that the same thing was found in the research of Sudrajat (2018), namely, the problem of educators who neglected their duties and functions, such as not analyzing the questions given to students.

Proportionally, the HOTS items were arranged in a smaller number because each assessment tool had to be adjusted to the processing time, so that there were no major obstacles for students who worked on it. In addition, the teacher still has difficulty in compiling the HOTS questions. It is possibly because the teacher has not attended the HOTS question preparation training.

Based on interviews with the teachers of grade XI students, all of them pay attention to the validity of the content when compiling the items because the suitability of the material is the main benchmark. Then linked to the results of validation by experts, the validity of the content contained in each item has also been completed, but what needs to be improved in the assessment instrument is the Operational Verb or *Kata Kerja Operasional* (KKO) on indicators that do not match the form of questions that appear in the questions.

Furthermore, the questions that teachers usually use are questions that are already available in the textbook that has been provided for student, especially in daily tests. However, there were also items that the teacher modified according to the conditions and material presented in the classroom. Testing the assessment instrument before being used as a measuring tool for the success of the learning process must indeed be done so that the learning objectives can be achieved properly, testing in this study was carried out as research supporting data to determine the quality of the items compiled, which in previous research was conducted by [Hartuti and Handayani \(2019\)](#) with no testing, so they only know the results of the analysis. In their research, it was found that the implementation of the 2013 curriculum assessment, in general, was in accordance with the 2013 curriculum assessment standards by making HOTS questions from the daily tests, mid-semester tests, and end-semester tests were in accordance with the syllabus, lesson plans, teacher books, and 2013 curriculum standards.

CONCLUSION

Based on the data analysis and research findings, it is concluded that (1) the instrument used by elementary school teachers in East Jakarta has fulfilled the content validity, (2) the cognitive domain (LOTS & HOTS) of the instruments used by elementary school teachers is proportional to 81.25% LOTS and 18.75% HOTS, (3) the quality of the assessment carried out by elementary school teachers in East Jakarta is generally good, but there are three items, namely two items for school E and one item for school F, which are not in accordance with basic competencies, and (4) implementation of assessments carried out by elementary school teachers in East Jakarta has followed the assessment standards given by the government. In addition, it is suggested that regular training for teachers should be conducted to monitor their success, and it is necessary to carry out comprehensive further training at all levels of school and in all subject areas as a step to maximize the implementation of the demands of the curriculum in the learning process. Assistance for teachers in training and developing themselves to maximize the learning process, including the assessments carried out, can also support their success.

REFERENCES

- Amirono, A., & Daryanto, D. (2016). *Evaluasi dan penilaian pembelajaran Kurikulum 2013*. Pustaka Pelajar.
- Arikunto, S. (2011). *Prosedur penelitian: Suatu pendekatan praktik*. Rineka Cipta.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals* (B. S. Bloom (ed.)). Longman.
- Brookhart, S. M. (2010). *How to assess 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. *Jurnal Review Pembelajaran Matematika*, 1(1), 86–98. <https://doi.org/10.15642/jrpm.2016.1.1.86-98>
- Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communication and Technology*, 29(2), 75–91. <https://www.jstor.org/stable/30219811>
- 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. *Journal of Medives: Journal of Mathematics Education IKIP Veteran Semarang*, 3(1), 55–63. <https://doi.org/10.31331/medivesveteran.v3i1.698>
- Khaerudin, K. (2015). Kualitas instrumen tes hasil belajar. *Jurnal Ilmiah Madaniyah*, 5(2), 212–235. <https://journal.stitpemelang.ac.id/index.php/madaniyah/article/view/26>
- Kurniasih, I., & Berlin, S. (2016). *Revisi kurikulum 2013: Implementasi dan konsep penerapan*. Kata Pena.
- Maolani, R. A., & Cahyana, U. (2015). *Metodologi penelitian pendidikan*. Rajawali Pers.
- 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 Sempoa: Seminar Nasional, Pameran Alat Peraga, Dan Olimpiade Matematika 1 2015*, 71–78. <https://publikasiilmiah.ums.ac.id/handle/11617/6132>
- Mulyasa, M. (2017). *Pengembangan dan implementasi Kurikulum 2013*. PT Remaja Rosdakarya.
- Nugroho, R. A. (2018). *HOTS: Higher order thinking skill*. PT Gramedia Widiasarana Indonesia.
- Nurani, H., Artharina, F. P., & Kiswoyo, K. (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://ejournal.undiksha.ac.id/index.php/IJERR/article/view/17625>
- Pi'i, P. (2016). Mengembangkan pembelajaran dan penilaian berpikir tingkat tinggi pada mata pelajaran Sejarah SMA. *Sejarah Dan Budaya: Jurnal Sejarah, Budaya, Dan Pengajarannya*, 10(2), 197–208. <https://doi.org/10.17977/um020v10i22016p197>
- Prasetya, T. I. (2012). Meningkatkan keterampilan menyusun instrumen hasil belajar berbasis modul interaktif bagi guru-guru IPA SMP N Kota Magelang. *Journal of Research and Educational Research Evaluation*, 1(2), 106–112. <https://journal.unnes.ac.id/sju/index.php/jere/article/view/873>
- Regulation of the Minister of Education and Culture No. 104 of 2014 on the Learning Outcome Assessment by Educators in Primary and Secondary Education Levels, (2014).
- Regulation of the Minister of Education and Culture No. 20 of 2017 on the Team of Performance Assessor in the Ministry of Education and Culture, (2017).
- Regulation of the Minister of Education and Culture No. 21 of 2016 on the Content Standard of Primary and Secondary Education, (2016).
- Regulation of the Minister of Education and Culture No. 23 of 2016 on Educational Assessment Standard, (2016).

- Samosir, A., Hasruddin, H., & Dongoran, H. (2019). Analisis kuantitas dan kualitas pertanyaan guru Biologi dan siswa materi Sistem Eksresi. *Jurnal Pelita Pendidikan*, 7(1), 9–15. <https://doi.org/10.24114/jpp.v7i1.10523>
- Setiadi, H. (2016). Pelaksanaan penilaian pada Kurikulum 2013. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 20(2), 166–178. <https://doi.org/10.21831/pep.v20i2.7173>
- 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.
- Situmorang, J. (2018). *Higher order thinking skills: Pengembangan keterampilan berfikir tingkat tinggi*. MDP Media.
- Sudrajat, H. (2018). *Analisis alat evaluasi pada mata pelajaran IPA di kelas V Sekolah Dasar Negeri 3 Nyerot Kecamatan Jonggat Lombok Tengah* [Graduate thesis, Universitas Maulana Malik Ibrahim, Malang]. <http://etheses.uin-malang.ac.id/11101/1/15761011.pdf>
- Suryanto, B., & Sutinah, S. (2011). *Metode penelitian sosial*. Kencana Media Group.
- Wachyudi, I., Sukestiyarno, S., & 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. <https://journal.unnes.ac.id/sju/index.php/jere/article/view/6928>
- Wardhani, D. F., & Putra, A. P. (2016). Pengembangan instrumen tes standar kognitif pada mata pelajaran IPA kelas 7 SMP di Kabupaten Banjar. *Proceeding Biology Education Conference*, 75–82. <https://jurnal.uns.ac.id/prosbi/article/view/5658>
- Wijayanto, P. A., Allifah, A., & Amirrudin, A. (2016). Evaluasi kualitas instrumen tes dalam pembelajaran Geografi di MAN 2 Kota Batu. *Jurnal Geografi*, 13(2), 101–113. <https://journal.unnes.ac.id/nju/index.php/JG/article/view/7969/5523>
- 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>