



## DIFFERENTIATION OF TEST ITEMS BETWEEN THE HIGH SCHOOL BIOLOGY OLIMPIAD IN NORTH KAYONG AND THE NATIONAL SCIENCE OLIMPIAD

R. Marlina\*<sup>1</sup>, H. Puspaningrum<sup>2</sup>, Hamdani<sup>3</sup>

<sup>1</sup>Department of Biology Education, Universitas Tanjungpura, Indonesia

<sup>2</sup>SMA N 2 Simpang Hilir, Indonesia

<sup>3</sup>Department of Physics Education, Universitas Tanjungpura, Indonesia

DOI: 10.15294/jpii.v6i2.10679

Accepted: March 28<sup>th</sup>, 2017. Approved: June 30<sup>th</sup>, 2017. Published: October 17<sup>th</sup>, 2017.

### ABSTRACT

The purpose of this research is to find out the differentiation between high school biology olympiad in the District of North Kayong and the National Biology Olympiad. The analysis is used to provide feedback to students regarding their knowledge in the forms of factual, conceptual, procedural, and metacognitive traits. This research is a descriptive study consisting of two phases: the stage of designing test items carried out by four Biology teachers who joined the group, Science Teachers Council, and the test tryout phase given to 33 high school students of class XI. This research resulted in the dimensions of knowledge which indicates that there are 79% (63 items) being in the dimension of the factual, as much as 15% (12 items) in the conceptual, as much as 6% (5 items) in the procedural, while the metacognitive dimension is 0%. The question package which was given in the preliminary phase test was 5% considered difficult, while in the final stage such difficulty was not found (0%). Therefore, it is concluded that the question items need to be revised because they have differentiation between high school biology olympiad in the District of North Kayong and the National Biology Olympiad.

© 2017 Science Education Study Program FMIPA UNNES Semarang

**Keywords:** knowledge dimension; bloom's taxonomy; difficulty index

### INTRODUCTION

The National Science Olympiad (NSO) is a science competition for students in elementary, junior and senior high schools, held by the Ministry of Education and Culture. This activity started in 2002. The NSO has become a medium for selecting Indonesian students with interests and competencies in science at the national level, who then can advance to the international level. Based on the results of interviews with the head of the Subject Teachers Council, in the implementation of coaching/guidance of candidates for high school biology olympiad held in schools spread in North Kayong District, 50% of the test items used comes from the previous years. The other 50% comes from a collection of the items

made by teachers who usually took them from the test questions of the school exam. This issue has not highlighted the distinctive traits of thinking for high school students. In fact, according to Khasanah et al. (2017), biology science requires distinctive traits of thinking. The problems are usually related to the uneven distribution of the material and also the unavailability of the key answers and the way to answer the questions

The test items made by science teachers in Indonesia still need a lot of improvement. This is evident from the unpreparedness of Indonesian students to compete globally. In addition, in various international standardized tests which have been taken by learners from Indonesia, the results are not satisfactory. Implementation of the 2013 curriculum is an attempt by Indonesia to catch up with other countries in the field of education, especially science (Pratiwi et al., 2016).

\*Address Correspondence:  
E-mail: renitahak@yahoo.com

The science lessons in the 2013 curriculum are conducted in an integrated way to provide opportunities for students to develop thinking skills, procedural skills, and scientific attitudes (Rosana et al., 2017). The preparation of test items for the Olympiad are undertaken to identify teachers' ability to prepare test items in accordance with the 2013 curriculum and revised Bloom's taxonomy. The teacher is a decisive component in the success of the teaching and learning process. According to Barinto (2012) teachers are the spearheads that directly relate to learners as subjects and objects of instruction. The competencies that must be possessed by a teacher include pedagogic competency, such as skills in preparing test items that match the dimension of knowledge and pay attention to the level in revised Bloom's taxonomy. In addition, teachers should also have skills in doing instructional evaluation. Instructional evaluation is a step to determine the characteristics of student achievement (Sasongko, 2016). One of the ways is with the activity of item analysis. This activity aims to measure whether or not the items that have been made and used belong to the good category.

The results of the item analysis can be used to revise the material to be measured and provide information to learners about the limits of knowledge they have. According to Anderson et al. (2001), the dimension of knowledge consists of four types: (1) factual knowledge; (2) conceptual knowledge; (3) procedural knowledge; and (4) metacognitive knowledge.

Factual knowledge covers the basic things that learners must know. Conceptual knowledge includes both explicit and implicit theories in different cognitive psychology. Procedural knowledge ranges from doing routine exercises to solving a new problem. Metacognitive ability helps learners to understand the material and solve problems encountered. Learners who use metacognitive strategies appropriately can be critical thinkers, problem solvers, and good decision makers rather than learners who do not use metacognitive strategies. Metacognitive thinking strategies are rarely applied in instruction (Septiyana et al., 2013).

Metacognitive thinking strategies can be applied by providing biology questions that require them to think critically. Critical thinking is an important skill to develop (Cahyarini et al., 2016). The test items that are capable of requiring students to think critically are the ones made by the National Science Olympiad (NSO) because NSO test items apply Higher Order Thinking (HOT). According Widodo & Kadarwati (2013), the app-

lication of HOT is capable of increasing interaction between students, and between students and teachers. Students will be more motivated to ask questions, express ideas and dare to solve a difficult question and use it in the national exam and or the National Science Olympiad (NSO) test.

National Science Olympiad (NSO) in the subject of biology at the district level is conducted annually. The test items are at the level of C4 (analyzing), C5 (evaluating), and the highest being C6 (creating). However, from results of pre-research conducted in one high school in North Kayong (February 2016), it was found that the items used in the early selection stage still used the cognitive domain at the lowest level of C1 which means at the remember level and C2 at understand level. Therefore, a study was conducted to identify the differentiation between the test items used in the preliminary selection process at the district level, especially in Kayong Utara District and the NSO test items. The differentiation was examined from 2 dimensions, i.e. the dimensions of knowledge and the item difficulty index.

## METHODS

This research is a descriptive study. It consists of two phases namely the phase of preparing the items and the test phase. The item preparation phase was carried out in North Kayong by six biology teachers who joined the Science Teachers Council in North Kayong. The data obtained from this phase 1 research were test items from Biology Olympiad of high school level in the form of multiple choice as many as 2 packages (Package 1 and Package 2), each package having 40 items. Package 1 was intended for use in the preliminary round and package 2 for the final round. Both packages were then processed descriptively and qualitatively by other Science Teachers Council participants (11 people) in order to map the items according to the revised Bloom's taxonomy.

The mapping aims to measure if the items are in the factual, conceptual, procedural, or metacognitive dimension. In addition, the calculation of the distribution was also done to determine the cognitive domain which ranges from remember (C1) to create (C6). Both packages of test items were validated by 2 lecturers of Biology Education at the Faculty of Teacher Training and Education (FKIP) Tanjungpura University. The type of validity used was content validity. This validity is chosen because this type of validity is often used in the assessment of learning outcomes.

In phase 2 the items were tried out on 33 students in Sungai Raya in the odd semester of the academic year of 2016/2017. The test tryout was conducted for 2 meetings in the same week. The first meeting was for items of the challenge round and the second meeting for items of the final round. Then, the calculation of reliability and item difficulty index were performed.

The multiple choice item reliability was calculated with Kuder Richardson (KR-20) using the following formula:

$$KR-20 = \frac{k}{k-1} \left[ 1 - \frac{\sum p(1-p)}{(SD)^2} \right]$$

where:

k : number of items  
(SD)<sup>2</sup> : variant

The difficulty index was calculated using the following formula.

$$P = \frac{B}{JS}$$

where:

P : Difficulty level  
B : Number of test takers who answer questions correctly  
JS : Total number of test takers

According to Arifin (2009) difficulty index is classified as follows:

P: 0.00 - 0.30 categorized as difficult  
P: 0.30 - 0.70 categorized as medium  
P: 0.70 - 1.00 categorized as easy

A number used to determine the level of difficulty of a test item is called a difficulty index. The difficulty index is in the range of 0.00 to 1.00. The difficulty index is also known as Proportion (P). If P 0.00, the item is very difficult; if P = 1.0, then the item is too easy (Syamsudin, 2012).

## RESULTS AND DISCUSSION

The selection of outstanding learners in the field of biology is carried out at the beginning of each semester every year. This activity is the agenda of the Education Office as a step to improve the quality of education in the district. So one of the objectives of the NSO in the field of biology is to select outstanding students in the

field of mathematics, science, and technology.

The selection is conducted in different levels starting from school to the district/city, provincial, national, and international level. The selection chart is shown in Figure 1 below. Based on Figure 1, school-level selection is the gateway for each learner to move on to the next stage. Each test item should have its level of difficulty measured and its knowledge dimensions reviewed.



Figure 1. NSO Stages

The theoretical characteristics of Biology Olympiad should only from the theory learned in high school or equivalent. However, to be able to answer the questions with good analysis, learners need insight, accuracy, and way of thinking. Thinking is all mental activity that helps formulate or solve problems to make a decision, understand, and seek answers (Husamah, 2015). This can be obtained from guidance and coaching with good quality questions. Activity in the form of coaching and mentoring of NSO is obviously needed by learners. Some learners complain about the lack of knowledge, understanding and solving NSO test items in Biology.

The number of test items successfully prepared by 4 Biology teachers who are members of the Science Teachers Council of North Kayong District amounted to 80 items in the form of multiple choice questions. The study on the quality of the items from the aspects of knowledge is shown in Figure 2 below.

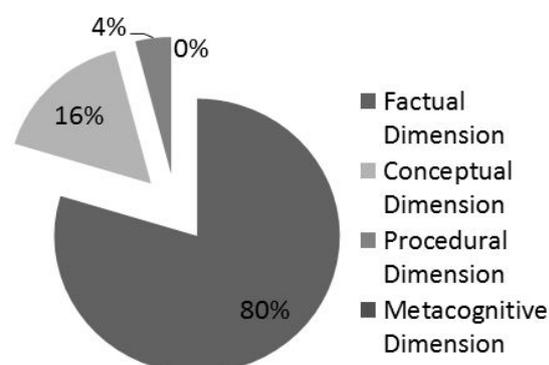


Figure 2. Distribution of Knowledge Dimension

Figure 2 shows that there are 63 items (80%) are in the factual dimension, 13 items (16%) on the conceptual dimension and 4 items (4%) on procedural dimension. While the meta-cognitive dimension is not found (0%). Based on the results of item compatibility analysis with the biology materials, 100% of the questions are in accordance with the material content outline tested in the selection of biology olympiad. Any types of test items that are prepared be it for daily test, school exam, and national exam, must be made based on the content outline and developed in accordance with the indicator of the item.

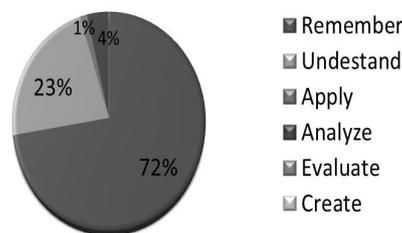
The problem is that the indicator of the Olympiad test items downloaded from the internet does not display the contents in detail. If the content outline of the Olympiad material does not appear in in test item, it will lower the ability of the item to measure what it should from the test results. Thus it was decided by the teachers who prepared for the Olympiad test items to reformulate the indicators of the items. According Maiza (2013) a test item is good if it has been in accordance with the content outline that has been formulated. If the item is not in accordance with the formulated outline, then it cannot function properly.

Based on Bloom's revised study of the distribution of taxonomy in the cognitive domain, it is found that most (50%) of questions are in C1 and C2 levels. Based on the results of the analysis of the items, it was found that a question can be easily answered if the student can remember the material in the textbook. In line with that, Rusyati & Rustaman (2013) argued that to answer a question, first most students have to read, understand and memorize the material in it. This is because the question is the as that found in the textbook, so they only need to remember what they have read from the textbook or from listening to the teacher's explanations.

In addition, the selection of questions in levels C1 and C2 is due to several other reasons as follows: (A) Teachers do not have enough time to think of or prepare items that demand learners' high-level of thinking; (B) Lack of access to references, and limited internet access both at work and in residential areas; and (C) Adjusting to the conditions of learners in the school.

In general, the results of the study of the distribution of test items in the cognitive domain are shown in Figure 3. Figure 3 shows that the items made by biology teachers still focus on the learner's memorization and comprehension. This is in line with Ningsih's (2016) study which stated that there has not been an item made by teachers

to measure the aspects of evaluate (C5) and create (C6).



**Figure 3.** Distribution of Items Based on Levels of Bloom's Taxonomy on The Cognitive Domain

Creativity in preparing test items still needs to be improved. The interview with the head of the Science Teachers Council of North Kayong District shows that based on experience during teaching, Biology questions that appear in the school exam is on the dimensions of factual and conceptual so that teachers feel that they have the obligation to familiarize the students with the material by giving questions to be memorized, consequently learners are not accustomed to thinking critically in answering questions.

Since most of the questions prepared in the cognitive domain of C1 and C2 levels, it will cause learners to be less capable of making new things or solving a problem. Problem solving is a model that can be used to improve learning outcomes (Triyuni, 2016). But the percentage of test items in the form of problem solving is lower than the items that need memorizing. The reason why teachers prepare more items that require memorizing in the cognitive domain levels C1 and C2 is related to educational goals today. There are two important educational goals i.e., to develop memory and to encourage transfer process. The occurrence of the transfer process is a sign of the success of a learning process. Memory or retention is the ability of a learner to recall learning materials for some time after teaching with the same accuracy as when the learners are following the lesson. Transferability is the ability of a learner to use what he has learned to solve new problems, to answer new questions, or to facilitate the learning of new things (Mayner, 1996). In short, the ability of memory is defined as the ability of a student to remember what he has learnt, while the transfer ability requires a learner to be able to remember and also understand and use what he has learned (Bransford et al., 1999).

Teacher's references in preparing Olympiad test items is also a standard of the quality of the items made in addition to the NSO syllabus.

bus. The items to be prepared should refer to the instructional syllabus and should at least be adjusted to the learning objectives contained in the valid or standard lesson plans. But in the process of preparing a test item, most teachers rely solely on the textbook. It means that teachers use the learning syllabus only in designing lesson plans as a complementary teacher administration to be applied in classroom instruction. Teachers hardly ever use the syllabus and lesson plans as one of the main sources in preparing the Olympiad test items. The correlation between knowledge dimension and aspects of Bloom's taxonomy is shown in Figure 4.

Based on the correlation between knowledge dimension and Bloom's taxonomy, the results of this study show the correlation from the distribution of the NSO test items. In Table 1 illustrates of distribution test items based on revised Bloom's taxonomy and knowledge dimension.

**Table 1.** Distribution Test Items Based on Revised Bloom's Taxonomy and Knowledge Dimension

Revised Bloom's Taxonomy	Knowledge Dimension			
	A	B	C	D
1	58	0	0	0
2	5	13	0	0
3	0	0	1	0
4	0	0	3	0
5	0	0	0	0
6	0	0	0	0
Total	63	13	4	0

Knowledge Dimension:

- A = Factual
- B = Conceptual

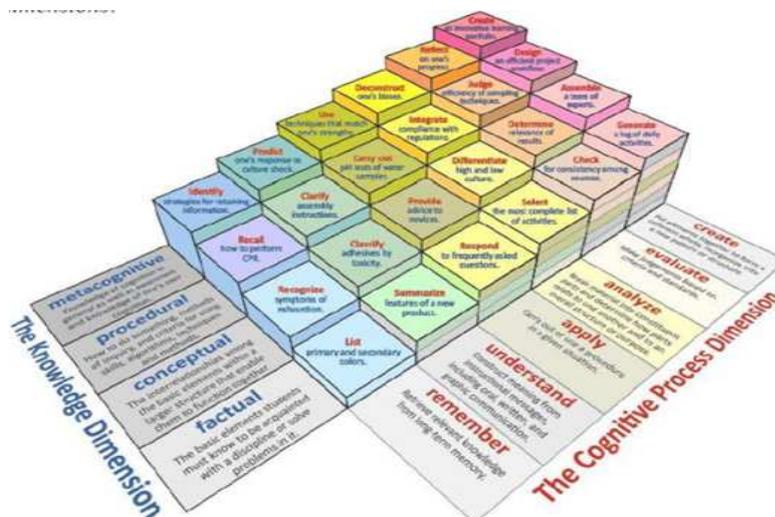
- C = Procedural
- D = Metacognitive

Bloom Taxonomy Revision:

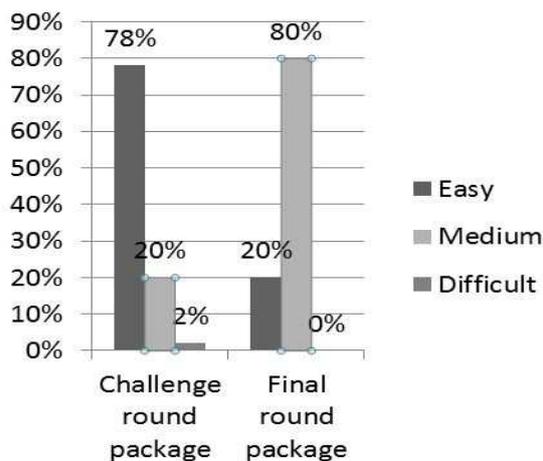
- 1 = Remember
- 2 = Understand
- 3 = Apply
- 4 = Analyze
- 5 = Evaluate
- 6 = Create

The distribution of test items prepared by teachers based on knowledge dimension and revised Bloom's taxonomy shows that most of the items are still at the level of remember and understand. The next items were tried out and the level of reliability calculated. The calculation of reliability aims to identify the level of accuracy and reliability of test scores. The calculation results have a range from 0 to 1. If the result is closer to the number 1, the reliability level of the item is higher. Based on the calculation of the data, the reliability score is R = 0.68. Based on the reliability criteria, the score of 0.68 belongs to the category of fair. This shows that the Olympiad test items prepared by teachers are quite reliable.

The items that have been examined from the knowledge dimension were then tried out to determine the reliability and the level of difficulty. In the try-out process, it was done 2 times with 2 days of lag time in the same week with different test packages. The results of the tryout analysis of the Biology olympiad test items can be seen in Figure 5.



**Figure 4.** Combination between Knowledge Dimension and Bloom's Taxonomy



**Figure 5.** Result of Analysis on Test Items Difficulty Level of Challenge Round Package and Final Round Package

The percentage of the prepared items should have a balance between easy, medium, and difficult. However, the results obtained from Figure 5 indicate the absence of balance of difficulty because in the challenge package, the most dominant are items with an easy category (78%) while the final package belongs to a medium category (80%). But in both packages, the percentage of difficult questions is very low (2% and 0% respectively). Test items with the results of difficulty analysis as shown in Figure 5 are not recommended for use in the Olympiad. This is in accordance with the opinion of Anastasia (2003) who stated that if the purpose of preparing test items is for a selection process, then it should have a medium or difficult level of difficulty. Items in the easy category should not be used. But should it be used, a few would be enough.

The item difficulty index of the data in Figure 5 reflects the low level of learners' experience to interact or work on a test item based on Higher Order Thinking Skill (HOTS). This type of HOTS items appears in the national exam and Trends in International Mathematics and Science Study (TIMSS). Based on the test tryout results, only 3 students who can correctly answer the questions in the difficult criteria. The results of interviews with 11 students who answered the items incorrectly show that the materials in the test items have not been learned before.

To this point, the test items they have done are those whose answer choices are definite and do not require analysis or high-level thinking processes. Being used to answering test items of memory and understanding alone has caused them to learn by memorizing in the absence of synthesis or analysis of the material learned. It will

obvious if the question is in the form of essay, where learners will have difficulty in giving an explanation of the material examined.

The findings of this research are significant for teachers and test developers. They must have the competence in analyzing the test material so that they can prepare the items up to the level of evaluate and create especially for the Olympiad test. This is because the Olympiad test items aim to select outstanding learners in the field of biology so it requires questions with higher difficulty level (Nurinda, 2014). Thus the Olympiad test items that have passed the phase of item analysis can be used as a basis in making the right decision. According to Boopathiraj & Chellamani (2013) to date, item analysis is an important phase in the development of tests or instruments. So it is expected that teachers and instrument developers can apply it in the final process of instrument preparation. For example in preparing Olympiad test items and the items that serve as a test of learning outcomes.

## CONCLUSION

There are 63 items (80%) in the factual dimension, 13 items (16%) in the conceptual dimension, and 4 items (4%) in the procedural dimension, while the metacognitive dimension is not found (0%). Among the 80 items prepared, 72% of them is at the remember level (C1) and 23% at the understand level (C2), only 1% at the apply level (C3), 4% at the analyze level (C4), while evaluate (C5) and create (C6) are not found (0%).

The compatibility of the olympiad test items prepared by the teachers with the indicators set in the olympiad content outline is 100%. The reliability of NSO items is the fair criterion ( $R = 0.68$ ). The item difficulty index shows that the most dominant items prepared in the challenge package belong to the easy category (78%). In the final package, the most dominant belongs to the medium category (80%). In both NSO packages, the lowest category is the difficult category (2% and 0%).

Therefore, it can be concluded that there is a differentiation between the test items used in the preliminary selection for the Olympiad in North Kayong and the items used in NSO (National Science Olympiad). There is a need for training on the preparation of items at the levels of C4, C5, and C6 in accordance with NSO test items so that it will improve student achievement in North Kayong District.

## REFERENCES

- Anastasia, U.S. (2003). *Psychological Testing*. New Jersey: Prentice-Hall.
- Anderson, L. W., Krathwohl, D. R., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Wittrock, M. (2001). *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy*. New York: Longman Publishing.
- Barinto. (2012). Hubungan Kompetensi Guru dan Supervisi Akademik dengan Kinerja Guru SMP Negeri Se-Kecamatan Percut Sei Tuan. *Jurnal Tabularasa PPS UNIMED*, 9(2), 201-214.
- Boopathiraj, C., & Chellamani, K. (2013). Analysis of Test Items on Difficulty Level and Discrimination Index in The Test For Research in Education. *International Journal of Social Science and Interdisciplinary Research*. 2(2). 128-134
- Bransford, J. D., Brown, A. L., & Cocking, R. (1999). *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: Committee on Developments in the Science of Learning.
- Cahyarini, A., Rahayu, S., & Yahmin. (2016). The Effect Of 5E Learning Cycle Instructional Model Using Socioscientific Issues (SSI) Learning Context on Students' Critical Thinking. *Jurnal Pendidikan IPA Indonesia*, 5(2), 222- 229.
- Husamah. (2015). Thinking Skills for Environmental Sustainability Perspective of New Students of Biology Education Department Through Blended Project Based Learning Model. *Jurnal Pendidikan IPA Indonesia*, 4(2), 110- 119.
- Khasanah, A.N., Sajidan, & Widoretno, S. (2017). Effectiveness of Critical Thinking Indicator-Based Module in Empowering Student's Learning Outcome in Respiratory System Study Material. *Jurnal Pendidikan IPA Indonesia*, 6(1), 187-195.
- Maiza, R. (2013). Analisis Butir Soal Tes Tertulis Tingkat SMP pada Olimpiade Fisika Se-Riau Tahun 2013. *Jurnal Fakultas Keguruan dan Ilmu Pendidikan Universitas Riau*, 1(1), 1-11.
- Mayner, R. E. (1996). *Teaching and Testing of Problem Solving*. 2<sup>nd</sup> ed. Oxford, UK: Pergamon.
- Ningsih, K. (2016). Kemampuan Guru MIPA Membuat Penilaian Pengetahuan. *Jurnal Pendidikan Matematika dan IPA*, 7(2), 44-54.
- Nurinda, S., Ely, R., & Saiful R. (2014). Analisis Butir Soal Olimpiade Biologi SMA Tingkat Kabupaten/Kota Tahun 2013. *Unnes Journal of Biology Education*, 3(1), 77-84
- Pratiwi, Y.N., Rahayu, S, & Fajaroh, F. (2016). Socioscientific Issues (SSI) in Reaction Rates Topic and Its Effect on The Critical Thinking Skills of High School Students. *Jurnal Pendidikan IPA Indonesia*, 5(2), 164-170.
- Rosana, D., Kadarisman, A., Maryanto, A., & Sugiharsono. (2017). The Evaluation of Science Learning Program, Technology and Society Application of Audio Bio Harmonic System with Solar Energy to Improve Crop Productivity. *Jurnal Pendidikan IPA Indonesia*. 6(1), 63-70.
- Rusyati, L., Rustaman, N., & Saefudin. (2013). Pengembangan Soal Pilihan Ganda Berpikir Kritis Inch dan Profil Pencapaiannya di SMA Negeri Kota Bandung pada Tema Penyakit Manusia. *Jurnal Pengajaran MIPA*, 18(1), 124-134.
- Sasongko, P. (2016). Validation of Science Achievement Test with The Rasch Model. *Jurnal Pendidikan IPA Indonesia*, 5(2), 268-277.
- Septiyana, K., Andreas, P.B.P., & Wulan, C. (2013). Jurnal Belajar sebagai Strategi Berpikir Metakognitif pada Pembelajaran Sistem Imunitas. *Unnes Journal of Biology Education*, 2(1), 56-69.
- Syamsudin. (2012). Pengukuran Daya Pembeda, Taraf Kesukaran, dan Pola Jawaban Tes (Analisis Butir Soal). *Jurnal Ilmu Tarbiyah*. 1(2), 211-219.
- Triyuni. (2016). The Influence of Science Learning Set Using Csientific Approach and Problem Solving Model on Learning Outcomes of Junior High School Students in The Subject of Heat and Temperature. *Jurnal Pendidikan IPA Indonesia*, 5(2), 177- 185.
- Widodo, T. & Sri. K. (2013). Higher Order Thinking Berbasis Pemecahan Masalah untuk Meningkatkan Hasil Belajar Berorientasi Pembentukan Karakter Siswa. *Jurnal Cakrawala Pendidikan*, 32(1), 161 -171.