# Development of Critical Thinking Skills and Argumentation Skills Assessment Instruments Towards 21st-Century Skills Based on Non-Routine Problems

Rusmini Rusmini[1,*] Suyono Suyono[2] Rudiana Agustini[3]

[1,2,3]*Department of Chemistry, Universitas Negeri Surabaya, Indonesia*
**Corresponding author E-mail: rusmini@unesa.ac.id*

**ABSTRACT**
This is development research that aims to produce a set of instruments for assessing student teachers' critical thinking and argumentation skills on chemistry. The instruments were developed through the stages of conceptualization, development of tests, and trials. Results indicate that the instrument was tested for theoretical validity by 3 expert chemists and empirical validity by 47 student teachers from Chemistry Education program. Using SPSS 23 test results data were analyzed for empirical validity with Pearson correlation and test reliability with Cronbach alpha. The results showed that assessment instruments of the critical thinking and argumentation skills developed were considered very valid theoretically and based on the results of respondents' responses. The assessment instruments of argumentation skills are all valid and have low reliability. The results of the self-observation questionnaire show that in general students were aware of their inability to solve critical thinking and argumentation skills questions, so they are challenged to think hard in solving answers and are motivated to learn further.

***Keywords**: Critical thinking skills, Argumentation, Non-routine problems.*

## 1. INTRODUCTION

The 21st century is the era of globalization. In this era, many competencies and skills need to be grown and developed. According to the Partnership for 21st Century Skills, there are 21st-century skills called 4C (critical thinking and problem solving, collaboration, creativity, communication) [1]. Critical thinking is a manifestation of higher order thinking because critical thinking skills can be seen as students' thinking abilities to compare two or more information they have [2]. The purpose of critical thinking is to provide case evidence, interpret what happened, and solve problems by involving logic, reasoning, and decision-making processes [3] [4]. Critical thinking skills are part of the life skills that students must possess, especially in developing reasoning, communication, and problem-solving faced by students in everyday life [5].

These critical thinking skills and communication skills can be possessed by students if they are trained in argumentation skills [6]. Therefore, argumentation skills are one of the thinking skills that must be mastered by students. Argumentation skills are activities that facilitate the understanding of cognitive activities in building scientific knowledge. Critical thinking skills and argumentation influence each other. In critical thinking, students are trained to reconstruct and evaluate arguments [7]. The analysis focuses on identifying assumptions, reasons, and claims and examining how they interact to form arguments [8]. Students' lack of critical thinking can be seen from their inability to provide arguments correctly, provide less logical assumptions, and provide little evaluation based on relevant facts [9]. Critical thinking skills are part of analytical skills or higher-order thinking skills. These skills include the ability to analyze arguments [10]. Increasing students' argumentation skill scores will increase critical thinking scores [11].

Graduate competency standards include attitudes, knowledge, and skills. For undergraduate students must master higher-order thinking skills [12]. For this reason, an instrument is needed to measure students' ability in higher-order thinking, in this case, critical thinking and argumentation skills that are valid and reliable. A good quality test if it meets the criteria is valid and reliable [13]. Critical thinking includes components of learning outcomes

in the cognitive aspect of the higher-order thinking category, so critical thinking can be measured using a test question [14] [15]. In the future, the assessment instrument of critical thinking and argumentation skills can be used to test students' abilities.

## 2. METHODS

The research design, which consists of four stages, namely: 1) conceptualization, 2) test development, 3) test trials and 4) testing [16]. In this study, it is limited to test trial stage. Conceptualization is understanding the concept to be measured, in this case, the concept being measured is the concept of chemistry in everyday life in the form of non-routine problems. Non-routine problems require creative thinking and the application of multiple strategies to understand the problem and do not have a direct way of dealing with questions. It takes the best way to find the best way to solve the problem. Non-routine problems tend to be more complex and more difficult than routine problems [17].

The second stage is the development of test questions. The test questions developed consisted of critical thinking and argumentation questions. Both are in the form of essay questions. There are 10 critical thinking questions and 6 questions for argumentation. At this stage, a review is carried out by chemist to review the questions that have been developed.

The third stage is a trial test. At this stage, content and construct validation was carried out by 3 experts in the field of chemistry. At this stage, content, and construct validity data were obtained from the theoretically developed assessment instrument. Assessment instruments that have been validated according to expert validators are tested on students. The trial was conducted on 47 Chemistry Education students at 7th semester to obtain empirical validity and reliability data. At this stage, student response questionnaires were also distributed to provide opinions on the assessment instruments developed. Theoretical validity was measured using a validation sheet to assess the feasibility of questions from 3 experts. Questions that are not valid according to experts are not continued at the trial stage. Furthermore, the assessment results are averaged, and the value is converted according to the criteria in Table 1.

**Table 1.** Level of validation criteria [18]

| Percentage | Criteria |
|---|---|
| 1.00 – 1.75 | Not feasible |
| 1.76 – 2.50 | Feasible enough |
| 2.51 – 3.25 | Feasible |
| 3.26 – 4.00 | Veri feasible |

**Table 2.** Criteria for degree of reliability

| Reliability Percentage (r11) | Criteria |
|---|---|
| 1.00 < r11 ≤ 0.20 | Very low |
| 0.20 < r11 ≤ 0.40 | Low |
| 0.40 < r11 ≤ 0.60 | Medium |
| 0.60 < r11 ≤ 0.80 | High |
| 0.80 < r11 ≤ 1.00 | Very high |

Empirical validity and reliability were analyzed by SPSS 23. Validity was tested by Pearson Correlation. A data is said to be valid if $r_{count} > r_{table}$ or if the value of sig. < 0.05. While the basis for decision making in the reliability test is If the Cronbach's Alpha value > $r_{table}$ then the questionnaire is declared reliable; and If the value of Cronbach's Alpha < $r_{table}$, the questionnaire is declared unreliable. However, if Cronbach's Alpha values are all <0.6, the basis for making decisions is reliability compared to the $r_{Table}$ value. With N of 47, the rTable value is 0.288 for a significant level of 5%. The criteria for the degree of reliability (r11) according to Guilford are shown in Table 2.

At the time of the trial, student response questionnaires were also distributed. The response questionnaire has 2 parts, namely the response to the assessment instrument and the student's self-observation response to critical thinking and argumentation skills. Student responses to assessment instruments on a scale of 1-4 then entered the criteria as shown in Table 1. While the student self-observation response questionnaire with a yes-no answer was analyzed using the Guttman scale. The results of the percentage of student responses are included in the criteria as in Table 3. The number of yes-no answers was calculated as a percentage to describe the student's condition. The results of the assessment were made a percentage with the formula:

$$percentage\ (\%) = \frac{\sum obtained\ score}{maximum\ score} X\ 100$$

**Table 3.** Gutman scale

| Question | Answer Score | |
|---|---|---|
| | Yes | No |
| Positive | 1 | 0 |
| Negative | 0 | 1 |

## 3. Results and Discussion

The development of assessment instruments for critical thinking and argumentation skills is useful for measuring and evaluating students' abilities in critical thinking and argumentation. Argumentation thinking skills are needed to improve critical thinking skills [6] [11]. At the development stage, the researcher made 10 critical thinking questions and 6 arguments. After the review process was carried out by 3 experts, there was 1 question that was judged not to be a critical thinking question. The question is a question about signs of chemical reactions which are considered to be rote questions at the C1 level. The levels of thinking in Bloom's Taxonomy are categorized into two, namely: low-level thinking (remembering, understanding, and applying) and higher-order thinking (analyzing, evaluating, and creating) [19] [20]. Critical thinking can be measured using the cognitive level of taxonomy bloom, namely on higher order thinking skills (analyzing, evaluating, and creating) [21] [22] [23]. This question was omitted and not submitted for validation assessment at a later stage, so there are only 9 validated questions. Meanwhile, 6 argumentative questions were considered appropriate by the reviewer so that 6 questions were submitted to the next process.

### *3.1. Theoretical Validity*

The following is a Table of expert validation results for critical thinking questions and argumentation questions. Validation includes content and constructs validity. Table 4 is the result of theoretical validation by experts on the quality of critical thinking skills and argumentation skills that were developed. The value listed is the average value of the three validators. The results of the expert's assessment were converted to Table 1 to assess the feasibility.

Based on the data in Table 4, it can be seen that all questions about critical thinking skills and argumentation skills are very feasible both in terms of content validity and construct validity (language and graphics). Critical thinking skills and argumentation skills that have been developed show that phenomena and questions are in accordance with critical thinking indicators (analysis, interpretation, and inference) as well as argumentation indicators (claims, data, warrants, backing, qualifications, and rebuttal). There is also the suitability of the phenomenon with the criteria for non-routine questions, the suitability of the answers to the questions on the question grid, and the suitability of the answers to the questions on the grid. For the linguistic criteria, it shows that the writing of the assessment instrument uses good and correct Indonesian. The terms used are precise and easy to understand and the language used is short and clear. Graphic criteria include accuracy in choosing the type and size of the font used, which can be read easily, and the suitability of the layout of text and images/Tables/graphic.

**Table 4**. Expert validation results for critical thinking skills and argumentation skills questions on content criteria and construct criteria

| Rated aspect | Critical Thinking | | | | | | | | | Argumentation Skills | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 1 | 2 | 3 | 4 | 5 | 6 |
| Average score | 3,6 | 3,7 | 3,7 | 4,0 | 3,7 | 3,9 | 3,7 | 3,5 | 4,0 | 3,8 | 3,8 | 3,9 | 3,8 | 3,8 | 3,9 |
| Criteria | VF | VF | VF | VF | VF | VF | VF | VF | VF | VF | VF | VF | VF | VF | VF |

Information. VF = very feasible

### *3.2. Empirical Validity and Reliability*

Data for the results of empirical validity and reliability of the instrumented assessment were tested using SPSS 23. Empirical validity was analyzed by Pearson correlation. To be valid or invalid compared the calculated r with the $r_{Table}$ and the sig value from the Pearson correlation Table with a significant level of 5%. Empirical reliability is used by Cronbach's Alpha.

#### *3.2.1 Empirical Validity and Reliability of Critical Thinking Skills*

The results of the validity of the critical thinking skills assessment instrument are presented in Table 5.

**Table 5.** The results of the validation of the critical thinking skills assessment instrument based on the results of the Pearson correlation

| Question number | Comparing R Values | | | Viewing Sig. Value | | |
|---|---|---|---|---|---|---|
| | $r_{count}$ | $r_{Table}$ | criteria | Sig. value | | criteria |
| 1 | 0.216 | 0.288 | invalid | 0.145 | 0.05 | invalid |
| 2 | 0.574 | 0.288 | valid | 0 | 0.05 | Valid |
| 3 | 0.48 | 0.288 | valid | 0.001 | 0.05 | Valid |
| 4 | 0.179 | 0.288 | invalid | 0.229 | 0.05 | invalid |
| 5 | 0.449 | 0.288 | valid | 0.002 | 0.05 | Valid |
| 6 | 0.58 | 0.288 | valid | 0 | 0.05 | Valid |
| 7 | 0.333 | 0.288 | valid | 0.022 | 0.05 | Valid |
| 8 | 0.561 | 0.288 | valid | 0 | 0.05 | Valid |
| 9 | 0.252 | 0.288 | Invalid | 0.088 | 0.05 | invalid |

Based on Table 5, item numbers 1,4, and 9 are invalid because they have a rcount < rTable and a sig value. > 0.05. So that, the data from these items are not used for the Cronbach Alpha reliability test. For the Cronbach Alpha reliability test this time using only valid items, namely item number 2,3,5,6,7,8 with a total of 6 items.

**Table 6**. Reliability Statistics

| Cronbach's Alpha | N of Items |
|---|---|
| .533 | 6 |

**Table 7.** Cronbach's alpha if item deleted

| Item | X1 | X2 | X3 | X4 | X5 | X6 |
|---|---|---|---|---|---|---|
| Cronbach's Alpha if Item Deleted | .464 | .568 | .476 | .471 | .524 | .412 |

In Table 6, the Cronbach Alpha value of 0.533 is obtained where the value is lower than 0.6 (not reliable). However, if all the items in the Item-Total Statistics Table are less than 0.60 then $r_{Table}$ is used to see the results of the reliability test decisions that we do. The $r_{Table}$ value with the total value of N = 47 with a 5% significance of 0.288. From the output in Table 8, the values in the column "Cronbach's Alpha if Item Deleted" are all <0.6. So that the basis for decision making is compared with the $r_{Table}$ value. If the value of Cronbach's Alpha > $r_{Table}$ then the questionnaire is declared reliable. If the value of Cronbach's Alpha < $r_{Table}$ then the questionnaire is declared unreliable [24]. It can be concluded that the items above are reliable. The average value of Cronbach's alpha is 0.4858. If interpreted into Table 2, then the degree of reliability (r11) according to Guilford is included in the category of moderate criteria.

### *3.2.2. Validity and Empirical Reliability of Argumentation Skills*

The results of the theoretical validity of the argumentation skills assessment instrument are based on the sig value and the $t_{count}$ value of the Pearson correlation. The results are presented in Table 8. With 47 respondents at a significant level of 5%, the $r_{Table}$ value is 0.288.

**Table 8.** Validation results of argumentation skill assessment instrument based on the results of the Pearson correlation

| Question number | Comparing R Values | | | Viewing Sig. Value | | |
|---|---|---|---|---|---|---|
| | $r_{count}$ | $r_{Table}$ | criteria | Sig. value | | criteria |
| 1 | 0.511 | 0.288 | valid | 0 | 0.05 | valid |
| 2 | 0.617 | 0.288 | valid | 0 | 0.05 | valid |
| 3 | 0.580 | 0.288 | valid | 0 | 0.05 | valid |
| 4 | 0.508 | 0.288 | valid | 0 | 0.05 | valid |
| 5 | 0.334 | 0.288 | valid | 0.020 | 0.05 | valid |
| 6 | 0.457 | 0.288 | valid | 0.001 | 0.05 | valid |

Based on the data in Table 8 shows that all items are valid, then all data on these items can be used for Cronbach's Alpha reliability test. In Table 9 the Cronbach Alpha reliability shows the Cronbach Alpha value of 0.419 where the value is lower than 0.6 (not reliable). However, if all the items in the Item-Total Statistics Table are less than 0.60 then the $r_{Table}$ was used to see the results of the reliability test decisions that we do. The basis for decision making in reliability testing is as follows: If the value of Cronbach's Alpha > $r_{Table}$ then the questionnaire is declared reliable. If the value of Cronbach's Alpha < $r_{Table}$ then the questionnaire is declared unreliable [24].

From the output in Table 10, the values in the column "Cronbach's Alpha if Item Deleted" are all <0.6. So that the basis for decision making is compared with the $r_{Table}$ value. The comparison results obtained are all item values in the column "Cronbach's Alpha if Item Deleted" have a value > $r_{Table}$ (0.288). It can be concluded that the questions above are reliable. The average value of Cronbach's alpha is 0.372. If interpreted to Table 2, the degree of reliability (r11) according to Guilford is included in the category of low criteria.

**Table 9.** Reliability Statistics

| Cronbach's Alpha | N of Items |
|---|---|
| .419 | 6 |

**Table 10.** Cronbach's alpha if item deleted

| Item | X1 | X2 | X3 | X4 | X5 | X6 |
|---|---|---|---|---|---|---|
| Cronbach's Alpha if Item Deleted | .358 | .338 | .310 | .377 | .442 | .404 |

Non-routine problems are of questions whose solutions require broader and unusual thinking because the procedures are not as clear or not as same as the procedures learned in class. This problem requires students to think critically, increase understanding of concepts, develop reasoning, develop abstract thinking skills, and transfer their knowledge into unusual situations [25]. Students who are not used to solving questions in the form of non-routine problems will have difficulty working on these questions. Moreover, the problem is done independently at a certain time. To make students reach the stage of thinking, which is more complex, it is necessary to continue to train both individually and in groups.

### *3.3. Student Response*

*3.3.1. Analysis of The Results of Student Responses to Each Question is Presented in The Table*

The data in Tables 11 show how students respond to each question developed on the critical thinking and argumentation skill assessment instrument according to content and construct criteria**.**

**Table 11.** Student responses to assessment instrument critical thinking skills on content criteria and construct criteria

| Rated Aspect | Critical Thinking Skills | | | | | | | | | Argumentation Skills | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 1 | 2 | 3 | 4 | 5 | 6 |
| Average score | 3,8 | 3,7 | 3,8 | 3,7 | 3,7 | 3,8 | 3,7 | 3,8 | 3,8 | 3,8 | 3,8 | 3,7 | 3,8 | 3,8 | 3,8 |
| Criteria | VF | VF | VF | VF | VF | VF | VF | VF | VF | VF | VF | VF | VF | VF | VF |

Based on the results of student responses to the critical thinking and argumentation skills assessment instruments developed, all were in the very feasible category.

*3.3.2. Student Self-Observation Responses*

The student's self-observation response is used to find out how the students' thinking skills. The answer that is very interesting from the results of the response is that students realize that their critical thinking skills and argumentation skills are still lacking. Although most students have worked on critical thinking skills and argumentation skills, students realize that the questions developed are not easy to do. These questions can challenge thinking power and need to think hard to solve them. This awareness motivates students to study further.

Self-observation is part of the self-regulation process of performance. Self-regulation is the process by which a person activates his cognitive intelligence into academic skills by guiding himself based on self-created behavior and ideas to achieve a goal. Self-regulation includes process performance and self-monitoring. The process performance involves self-observation. Self-observation involves behavior and interests. Self-observation involves recording the frequency, intensity, or quality of behavior. Self-observation is very important in determining one's progress. Self-observation can lead to increased motivation because when people are aware of what they are doing, they can change their behavior [26]. By doing self-observation like this, students will be motivated to improve their abilities. High motivation can improve student learning processes and outcomes, interest, choosing tasks, efforts to learn difficult tasks, and patience [27].

## 4. CONCLUSION

The results showed that assessment instruments of the critical thinking and argumentation skills developed were considered very valid theoretically and based on the results of respondents' responses. Based on the empirical results there are 3 questions about critical thinking skills that were not valid and the reliability of the questions at the medium level. While the assessment instruments of argumentation skills are all valid and have low reliability. The results of the self-observation questionnaire show that in general students were aware of their inability to solve critical thinking and argumentation skills questions, so they are challenged to think hard in solving answers and are motivated to learn further. The results of this study still do not provide results as expected, namely a valid and high-level reliable assessment instrument. This provides an opportunity to further develop the two instruments. The use of non-routine problems in questions of critical thinking and arguments is very well used to be able to further develop students' thinking skills. This assessment instrument was tested on 7th-semester students but gave unsatisfactory results. This shows the importance since the early semesters even since elementary school to activate students' thinking skills.

## AUTHORS' CONTRIBUTION

Rusmini: developing the concept, creating timeline of writing articles, writing the manuscript. Suyono: reviewing the manuscript, adding data analysis. Rudiana Agustini: reviewing data analysis and presenting data.

**ACKNOWLEDGMENTS**

Our gratitude goes to the Chancellor of the Universitas Negeri Surabaya who has provided funding for this research through a doctoral dissertation grant in 2021.

**REFERENCES**

[1] Partnership for 21st Century Skills (P21), Framework for 21st century learning. Retrieved from P21 Partnership for 21st Century Skills, Framework for 21st Century Learning, 2015, Retrieved from *http://www.p21.org/our-work/p21-framework*

[2] D. R. Rizaldi, M. Makhrus, A. Doyan, Analisis Tingkat Kemampuan Berpikir Kritis Dengan Model Perubahan Konseptual Ditinjau dari Gaya Belajar Siswa, Jurnal Pendidikan Fisika dan Teknologi 5(1) (2019) 74-81. DOI: http://dx.doi.org/10.29303/jpft.v5i1.794

[3] A. Syawaludin, Gunarhadi, P. Rintayati P, Development of Augmented Reality Based Interactive Multimedia to Improve Critical Thinking Skills in Science Learning, International Journal of Instruction, 12(4) (2019) 331-344. DOI: https://doi.org/10.29333/iji.2019.12421a

[4] S. Arifin, P. Setyosari, C. Sa'dijah, D. Kuswandi, The Effect of Problem-Based Learning by Cognitive Style on Critical Thinking Skills and Students' Retention, Journal of Technology and Science Education, 10(2) (2020) 271-281. DOI: http://dx.doi.org/10.3926/jotse.790

[5] B. Harjo, B. Kartowagiran, A. Mahmudi, Development of Critical Thinking Skill Instruments on Mathematical Learning High School, International Journal of Instruction, 12(4) (2019) 149-166. DOI: https://doi.org/10.29333/iji.2019.12410a

[6] E. Roviati, A. Widodo, Kontribusi Argumentasi Ilmiah dalam Pengembangan Keterampilan Berpikir Kritis, Titian Ilmu: Jurnal Ilmiah Multi Sciences, 11(2) (2019) 56-66. DOI: https://doi.org/10.30599/jti.v11i2.454

[7] Irwanto, A. D. Saputro, E. Rohaeti, A. K. Prodjosantoso, Promoting Critical Thinking and Problem-Solving Skills of Preservice Elementary Teachers through Process-Oriented Guided-Inquiry Learning (POGIL), International Journal of Instruction, 11(4) (2018) 777-794. DOI: https://doi.org/10.12973/iji.2018.11449a

[8] J. E. Dowd, R. J. Thompson, L. A. Schiff, J. A. Reynolds, Understanding the Complex Relationship between Critical Thinking and Science Reasoning among Undergraduate Thesis Writers, CBE—Life Sciences Education, 17(4) (2018) 1–10. DOI: https://doi.org/10.1187/cbe.17-03-0052

[9] A. Muhlisin, H. Susilo, M. Amin, F. Rohman, Improving Critical Thinking Skills of College Students Through RMS Model for Learning Basic Concepts in Science, Asia-Pacific Forum on Science Learning and Teaching, 17(1) (2016) 1–24. Retrieved from https://www.eduhk.hk/apfslt/

[10] I. N. Suardana, I W Redhana, A. A. I. A. R. Sudiatmika, I. N. Selamat, Students' Critical Thinking Skills in Chemistry Learning Using Local Culture-Based 7E Learning Cycle Model, International Journal of Instruction, 11(2) (2018) 399-412. DOI: https://doi.org/10.12973/iji.2018.11227a

[11] A. Karadeniz, The Relationship Between Faculty of Education Students' Argumentation Skills and Critical Thinking, Creative Thinking and Problem-Solving Skills, TOJET: The Turkish Online Journal of Educational Technology, Special Issue for INTE, 2016. DOI:

[12] Kemendikbud, Permendikbud No 3 Tahun 2020 tentang Standar Nasional Perguruan Tinggi, Kementerian Pendidikan dan Kebudayaan Republik Indonesia, 2020.

[13] S. Arikunto, Dasar-Dasar Evaluasi Pendidikan, PT Bumi Aksara, 2009.

[14] Kartimi, Liliasari, Implementasi pengembangan alat ukur berpikir kritis pada konsep kesetimbangan kimia untuk Siswa SMA. Jurnal Scientiae Educatia, 1(2) (2012) 1–22. DOI: http://dx.doi.org/10.24235/sc.educatia.v1i2.505

[15] S. D. S. Pradana, Parno, S. K. Handayanto Pengembangan Tes Kemampuan Berpikir Kritis Pada Materi Optik Geometri Untuk Mahasiswa Fisika, Jurnal Penelitian dan Evaluasi Pendidikan, 21(1) (2017) 51–64. DOI: http://dx.doi.org/10.21831/pep.v21i1.13139

[16] J. Syahfitri, H. Firman, S. Redjeki, S. Srivati Development and Validation of Critical Thinking Disposition Test in Biology, International Journal of Instruction, 12(4) (2019) 381-392. DOI: https://doi.org/10.29333/iji.2019.12425a

[17] N. M. Murdiyani, Developing Non-routine Problems for Assessing Students' Mathematical Literacy, in: Proceedings of International Conference on Mathematics, Science and Education (ICMSE), IOP Conference Series: Journal of Physics Conference Series, 2018, pp. 1-6.

[18] R. B. Rudibyni, R. Perdana, E. Elisanti, Development of Problem-Solving-Based Knowledge Assessment Instrument in Electrochemistry, International Journal of Instruction, 13(4) (2020) 957-974. DOI: https://doi.org/10.29333/iji.2020.13458a

[19] C. P. Magas, L. D. Gruppen, M. Barrett, P. H. Dedhia, G. Sandhu, Intraoperative Questioning to Advance Higher-Order Thinking, American Journal of Surgery, 213(2) (2017) 222–226. DOI: https://doi.org/10.1016/j.amjsurg.2016.08.027

[20] L. Widiawati, S. Joyoatmojo, Sudiyanto, Higher Order Thinking Skills as Effect of Problem Based Learning in the 21st Century Learning, International Journal of Multicultural and Multireligious Understanding, 5(3) (2018) 96–105. Retrieved from https://ijmmu.com/index.php/ijmmu/article/view/223/139

[21] M. D. Saputra, S. Joyoatmojo, D. K. Wardani, K. B. Sangka, Developing Critical-Thinking Skills through the Collaboration of Jigsaw Model with Problem-Based Learning Model. International Journal of Instruction, 12(1) (2019) 1077-1094. DOI: https://doi.org/10.29333/iji.2019.12169a

[22] B. FitzPatrick, H. Schulz, Do curriculum outcomes and assessment activities in science encourage higher order thinking? Canadian Journal of Science, Mathematics and Technology Education, 15(2) (2015) 136–154. DOI: https://doi.org/10.1080/14926156.2015.1014074

[23] L. W. Anderson, D. R. Krathwohl D R, A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives, Longman, 2001.

[24] J. Widiyanto, SPSS for Windows Untuk Analisis Data Statistik dan Penelitian, BP-FKIP UMS, 2010.

[25] S. N. Mayangsari, L. T. Mahardhika, Scaffolding Pada Penyelesaian Soal Non Rutin Telescopic, Jurnal Ilmiah Edutic, 4(2) (2018) pp. 44-52. DOI: https://doi.org/10.21107/edutic.v4i2.3952

[26] B. Çetin, Academic Motivation and Self-Regulated Learning In Predicting Academic Achievement In College, Journal of International Education Research, 11(2) (2015) 95-106. DOI: https://doi.org/10.19030/jier.v11i2.9190

[27] B. J. Zimmerman, Goal Setting: A key proactive source of academic self-regulation, in D. H. Schunk, B. J. Zimmerman (Eds.), Motivation and Self-Regulated Learning Theory, Research, and Applications, Routledge Taylor &Francis Group), 2012.