



Development of Argumentation Skills Assessment Instruments on Buffer Solution Material

Fitria Rahmatul Ula¹, Suyono^{2*}

¹Fitria Rahmatul Ula, Chemistry Education Study Program, State University of Surabaya, Surabaya, Indonesia.

²Suyono, Chemistry Department, State University of Surabaya, Surabaya, Indonesia.

Received: July 21, 2023

Revised: September 4, 2023

Accepted: December 20, 2023

Published: December 31, 2023

Corresponding Author:

Suyono

suyono@unesa.ac.id

DOI: [10.29303/jppipa.v9i12.4760](https://doi.org/10.29303/jppipa.v9i12.4760)

© 2023 The Authors. This open access article is distributed under a (CC-BY License)



Abstract: This research aims to produce a qualified argumentation skill assessment according to the criteria of validity and reliability. This research is a development research with the ADDIE model whose implementation is limited to the pilot implementation and the evaluation formative evaluation. The subject of this research was an argumentation skills assesment in buffer solutions implemented in 18 students grade XI of SMA Trensains Tebuireng. The research instruments are review sheets, validation sheets, and argumentation skill test sheets. The results of content validity gain a mode with scores of 4 and 5 with valid and very valid criteria. In the construct aspect, gain a mode with a score of 5 with very valid criteria. The empirical validity result of the five questions in order is 0.791; 0.876; 0.961; 0.927; and 0.871. The calculated value appropriate with the criteria of rcalculate > rtable so that it can be declared valid. The value of reliability obtained 0.922 is greater than the rtable which is 0.468. With such results, the instrument can be declared reliable. Based on the results of the analysis, the argumentation skill assessment instrument is qualified to be used to measure students' argumentation skills on buffer solution material.

Keywords: Argumentation Skills; Assesment Instrument; Buffer Solution; Development; Reliability; Validity

Introduction

Research Background

Based on Permendikbud No. 34 tahun 2018 about content standards that in the learning process should integrating 21st century skills. 21st century skills called the 4Cs include critical thinking and problem solving, collaboration, creativity, and communication (Irvan et al., 2020). Argumentation skills are very important for students because argumentation skills include 2 important skills of the 21st century, that are critical thinking skills and communication skills (Devi et al., 2018).

Argumentation is an attempt to convince or prove the truth of a statement, opinion, attitude or belief, by being proven by facts, so as to be able to convince and prove that the opinion can be said to be true or not (Rahayu et al., 2020). Through his famous book entitled *The Uses of Argument*, Stephen Toulmin has made a

definition of Tolmin's argument known as Toulmin's Argument Pattern (TAP) consists of claim, data, warrant, backing, qualifier, and rebuttal (Abduh et al., 2019).

Research by Devi et al (2018) generate that students' argumentation skills is at level 1 and 2 argumentation. This statement shows that students' argumentation skills are still at a low-medium level (only telling claims or claims with data). The written argumentation skills of students in Sukabumi are still in the low category, this is shown by the highest percentage results at level 1, which is 38% and level 2 as much as 36% (Rahayu et al., 2020). The scientific argumentation skills of students in Payakumbuh are at a low level of ability and some are even very low (Permata Sari, 2018). The Indonesian students also cannot give the correct theory to support claims (Ain et al., 2018). Rahmawati & Suprpto (2019) stated that students are still not familiar to doing the types of argumentation skills questions and tend to be

How to Cite:

Ula, F. R., & Suyono. (2023). Development of Argumentation Skills Assessment Instruments on Buffer Solution Material. *Jurnal Penelitian Pendidikan IPA*, 9(12), 10792–10799. <https://doi.org/10.29303/jppipa.v9i12.4760>

more familiar to doing mathematical form questions only.

Chemistry is a branch of Natural Sciences (IPA) that studies the composition, structure, properties, changes in matter, and energy that accompany these changes. Chemistry includes all matter in nature (Amalia et al., 2019). One of chemistry concept is buffer solution. Argumentation skills can be applied to buffer solution material so that students can gain a deep understanding of these difficult topics. Through argumentation activities, students can understand concept acquisition process of knowledge being (Dwiretno & Setyarsih, 2018). Argumentation can be used to find out the student's level of understanding (Kamilahrohrawati, 2018). The low learning outcomes obtained by students show that there is still a low understanding of chemistry concepts that they have (Elyani et al., 2019). In the learning process, students are rarely given examples of questions with the type of scientific argumentation, students are mostly only given mathematical type questions, besides that, teachers lack of guidelines regarding how to evaluate and develop scientific argumentation skills in the classroom (Syerliana et al., 2018). Students must be able to state during argument presentations in accordance with standards and elements of scientific argumentation skills, such as claim, data evidence, and support (Jumadi et al., 2021). Argumentation skills are appropriate when applied to buffer solution materials included in KD 3.12. Where KD 3.12 is enriched with argumentation skills. One of the efforts that can be made to develop students' scientific argumentation skills is the availability of argumentation skills assessment instruments. Learning assessment is a process of obtaining information used to make decisions about students, curriculum and programs, and education policies (Nitko, as cited in Ambiyar & Panyahuti, 2020).

Assessment instruments as measuring instruments must satisfy certain criteria so that they can provide reliable information, which they must have good validity and reliability (Slamet & Wahyuningsih, 2022). Based on the description of the problem, the aim of this research is producing a feasible assessment instrument on the buffer solution material. The instrument is reviewed from the validity of both theoretical validity, empirical validity and reliability.

Theoretical Background

Assessment refers to all the information a teacher collects about his or her students in the classroom. This information can be obtained through formal tests in the form of essays and homework, as well as through informal tests in the form of observation and interaction (Arends, as cited in Sahidu et al., 2020). Assessments are

not only used for knowing learning outcomes however, can also be used as an important factor in the success of the student's learning process (Fidia et al., 2022). The quality of the assessment instruments is crucial for revealing the results and the learning process as its objectivity of the assessment outcomes, and its implementation also have a crucial role (Du et al., 2020). An important and integral component in teaching and learning activities are Assessment activities (Imania & Bariah, 2019). In educational activities, assessment refers to the process of measuring students' knowledge and abilities (Rafiqoh, 2020). Instruments are tools used to measure objects in order to fulfill academic criteria (Dachliyani, 2020).

A test declared as valid if it measures what is supposed to measure. So, in other words, the test is valid if it really measures what it is wanted to measure (Setiabudi et al., 2019). When applying a test's validity, a few factors need to be considered, which: suitability to the material to be tested; refers to the outcomes of an assessment or test administered to a group of people; concerning the degree of the validation terms high, medium, and low; and refers to the use of evaluation results. An assessment tool's validity can be understood in a number of ways, one of which is that it is determined by how accurately test findings or evaluation tools are interpreted for a certain group, not by the tool itself (Erlinawati & Muslimah, 2021). The term "content validity" describes how well a test's material relates with the curriculum. It assesses the elements as recommended by the curriculum (Rizky Hilaldy, 2021). An instrument that has undergone validity testing is capable of measuring what needs to be measured, and the outcomes of this empirical validity are grounded in empirical facts and indicate the degree of validity or accuracy of each item currently in use (Marselina et al., 2021).

Reliability is the quality of being consistent and trustworthy. A test will be reliable when it gives the same repeated result under the same conditions (Setiabudi et al., 2019). The purpose of reliability evaluation is to ascertain the stability of the measuring tools so that it maintains its stability when remeasuring (Azwar, 2018).

Method

This research is a type of development research. This research refers to the ADDIE model namely analysis, design, development, implementation, and evaluation. The implementation phase used only the pilot implementation and the evaluation conducted was a formative evaluation. The analysis stage consist of needs analysis, curriculum analysis, and student

analysis. The second stage conducted was design which included gathering supporting references, writing test questions, writing scoring rubrics, and making research instruments.

The third stage is development which includes the preparation of products in the form of assessment instruments. Before the product is implemented on students, a theoretical validity test is conducted by three validators, that are two Unesa chemistry lecturers and one high school chemistry teacher. The instruments used include content and construct validation sheets.

The types of research instruments used were review sheets, validation sheets, and argumentation skills test sheets. The review sheet contains suggestions and comments given by reviewers for product development. The validation sheet is used to assess the theoretical validity of the developed assessment instrument. The argumentation skill test sheets were tested on students to find out the empirical validity and reliability of the questions on the developed assessment instrument.

The data were analyzed descriptively through suggestions from the reviewers to improve the developed assessment instrument. Data from the theoretical validation results from the three validators were analyzed descriptively based on the Likert scale as shown in the following table 1.

Table 1. Likert scale score

| Response Value | Category |
|----------------|-----------------|
| 5 | Perfectly Valid |
| 4 | Valid |
| 3 | Sufficient |
| 2 | Less Valid |
| 1 | Not Valid |

Data obtained from experts are then analyzed using the mode or value that appears frequently. Based on the validator's assessment, the assessment instrument is declared valid if each question has a minimum Mo of 4 (Afni & Suyono, 2021).

Test results data were analyzed to determine the empirical validation and reliability of the developed assessment instrument. The validity test was obtained from the pearson product moment correlation test by correlating the score of each test item with the total test score. The validity test of each item can be stated whether valid or not based on a comparison between the Pearson correlation coefficient (rcount) and the Pearson table moment value (rtable). Items are stated to be valid if the value of rcount > rtable. The level of reliability is obtained from the results of the Cronbach alpha test (Yusup, 2018). Items are stated to be reliable if the value

of rcount > rtable. The scheme of research is shown in figure 1.

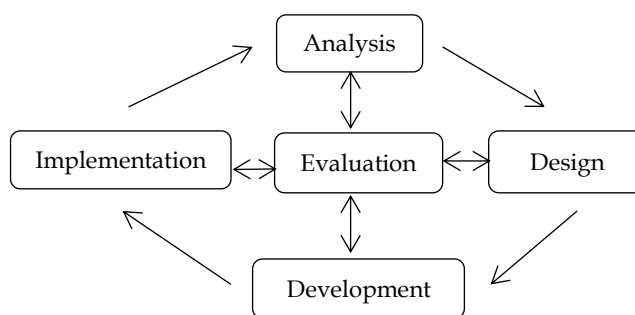


Figure 1. Scheme of Research

Result and Discussion

Analysis Results

The analysis phase consist of needs analysis, curriculum analysis, and student analysis. Needs analysis was conducted by field observations at schools and it was found that so far chemistry teachers at schools have not implemented learning with argumentation skills to determine students' understanding of a chemistry matter. In addition, the questions given by the teacher for both daily tests and end-of-semester assessments were in the form of essay and didn't yet apply indicators of argumentation skills. Curriculum analysis was conducted by examining Basic Competencies (KD) in formulating learning objectives so that the products are in accordance with the applicable curriculum. The curriculum used is the 2013 curriculum at KD 3.12 which explains the working principle, pH calculation, and the role of buffer solutions in living things. In this research, the students involved in the implementation phase were 18 high school students. Based on the results of observations and interviews conducted by researchers on May 9, 2023, in general, students took part in learning well. Students are quite active by asking and answering questions from the teacher. Indirectly, students are able to give an opinion which is one indicator of argumentation skills.

At the analysis stage, there is an evaluation of the argumentation skill indicators used, where previously only 5 indicators were used without rebuttal indicators. After being given directions by the advisor, a revision was made to add rebuttal to the indicator of argumentation skills used.

Design Results

At the design stage, the activities conducted include collecting supporting references, writing test questions, writing scoring rubrics, and preparing research instruments. Collecting supporting references was conducted by studying the literature related to the buffer

solution matter so that researchers can formulate a questions. The buffer solution matter includes the definition, working principle, and calculation of the pH of the buffer solution. The next stage is writing test questions and writing a scoring rubric. The argumentation skills questions is an essay form consist of 5 questions with 6 stages according to Toulmin's argumentation skill indicators. The scoring rubric of the test items consists of a description of each argumentation skill indicator and its score, where each indicator has a maximum score of 3. In addition, an assessment and score category is also added to determine students' argumentation skills. At the design stage, research instruments that are review sheets and validation sheets were also prepared. At the design stage, the product draft has been prepared and needs to be evaluated both in the question section and other complementary parts.

Development Results



Figure 1. Argumentation Skills Questions

At the development stage, the assessment instrument is prepared. The assessment instrument consists of a cover page, instrument description, argumentation skill indicators, examples of argumentation skill practice, argumentation skill questions and a scoring rubric. The display of the question is shown in figure 1.

The evaluation of the development stage was conducted by product validation. The assessment instrument was validated by three validators. The validation process produces value validation and suggestions for revision before being implemented on students. Validation is reviewed from the aspect of content and construct. The results of content validity can be seen in Table 2.

Table 2. Content Validity Results

| Num ber | Assessed Components | Validator | | | Mo de | Validity Criteria |
|------------|---|-----------|---|---|----------|----------------------|
| | | 1 | 2 | 3 | | |
| 1. | The questions presented have relevance to the material contained in the Basic Competency. | 5 | 4 | 5 | 5 | Perfectly Valid |
| 2. | The facts, concepts, and images presented are relevant to the material. | 4 | 4 | 5 | 4 | Valid |
| 3. | Content/topics have relevance to argumentation skills, that are: 1) Determine Claims 2) Show an evidence 3) Compile a warrant 4) Compile a backing 5) Compile a qualifier 6) Compile a rebuttal | 4 | 5 | 5 | 5 | Perfectly Valid |

Based on Table 2, the assessment instrument obtained a mode with a score of 4 and 5. Based on the validator's assessment, the assessment instrument was declared valid if each question had a minimum Mo of 4 (Afni & Suyono, 2021). The assessment instrument developed can be stated as valid and very valid criteria in terms of content aspects which include the suitability of the content/topics presented with the material contained in the Basic Competency and indicators of

argumentation skills. Furthermore, the results of construct validation can be seen in Table 3.

Table 2. Construct Validity Results

| Number | Assessed Components | Validator | | | Mode | Validity Criteria |
|--------|--|-----------|---|---|------|-------------------|
| | | 1 | 2 | 3 | | |
| 1. | There are directions for students to write a statement which is a claim that is structured in providing an assessment of a phenomenon (determining a claim). | 5 | 5 | 5 | 5 | Perfectly Valid |
| 2. | There are directions for students to write down some appropriate evidence to strengthen the claims that have been determined (showing an evidence). | 5 | 5 | 4 | 5 | Perfectly Valid |
| 3. | There are directions for students to formulate statements to explain the suitability of the evidence submitted in strengthening the claims that have been determined (compiling a warrants). | 5 | 5 | 5 | 5 | Perfectly Valid |
| 4. | There are directions for students to write statements/reinforcement based on the theory underlying claims or data (compiling a backing). | 5 | 5 | 4 | 5 | Perfectly Valid |
| 5. | There are directions for students to write statements regarding the relationship between claims and warrants (compiling a qualifiers). | 5 | 5 | 5 | 5 | Perfectly Valid |
| 6. | There are directions for students to write a statement and include the reasons proposed to say that a statement submitted is wrong (compiling a rebuttals). | 5 | 5 | 5 | 5 | Perfectly Valid |

Based on Table 3, the assessment instrument obtained a mode with a score of 5. Based on the validator's assessment, the assessment instrument was declared valid if each question had a minimum Mo of 4 (Afni & Suyono, 2021). The assessment instrument developed can be stated to be perfectly valid in terms of the construct aspect. Each component in construct validation represents an indicator of argumentation skills, so that it can be concluded that the assessment instrument developed contains directions for students to write claims, evidence, warrants, backing, qualifiers, and rebuttals sequentially.

At the development stage, product validation was conducted by three validators where there were suggestions from the validators for revised. The product revision results are shown in the Figure 2 and 3.

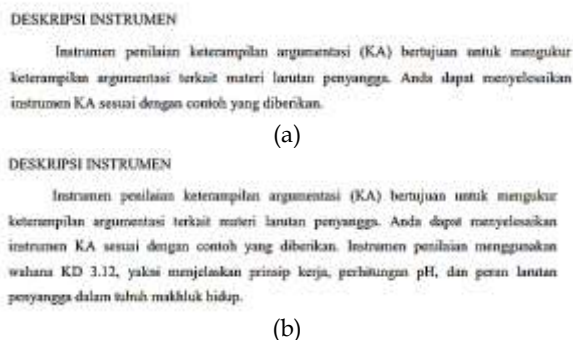


Figure 2. Instrument Description (a) Before Revision and (b) After Revision

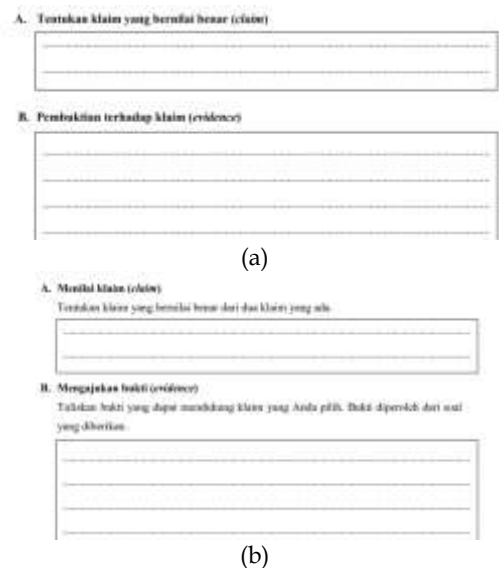


Figure 3. Question's Instructions (a) Before Revision and (b) After Revision

Implementation Results

At the pilot implementation stage, the researcher gave an instrument for assessing argumentation skills to 18 students. Students then work on the 5 questions that already exist by paying attention to the directions given in the instrument. In the evaluation of the implementation stage, the total score of the students was obtained after conduct the argumentation skills test. The total score of students is processed to get the results of empirical validity and reliability. The validity test was calculated using IBM SPSS with the Pearson product moment correlation test. The validity test uses Pearson Product Moment which is useful for determining the validity of the instruments used by researchers in

measuring and obtaining research data from respondents (Hidayati et al., 2023). The results of the validity of the questions can be seen in Table 4.

Table 4. Empirical Validity Test Results

| Number | r_{count} | Criteria |
|--------|-------------|----------|
| 1 | 0.791 | Valid |
| 2 | 0.876 | Valid |
| 3 | 0.961 | Valid |
| 4 | 0.927 | Valid |
| 5 | 0.871 | Valid |

The validity test of each item can be stated as valid or not based on a comparison between the Pearson correlation coefficient (r_{count}) and the Pearson table moment value (r_{table}). The r table value used for $n = 18$ with a significance level of 5% is 0.468. Based on Table 4 it can be seen that of the 5 questions fulfilling $r_{count} > r_{table}$, so that the 5 questions can be declared valid. Items are said to be valid if the value of $r_{count} > r_{table}$. A valid assessment instrument means that the instrument measures what should be measured, that is students' argumentation skills.

Reliability is calculated as a whole, that are 5 questions on the assessment instrument using the Cronbach alpha test. The level of reliability is obtained from the results of the Cronbach alpha test (Yusup, 2018). Alpha Cronbach was chosen as the reliability test because by doing the Cronbach's alpha test inconsistent indicators would be detected and Cronbach's Alpha is a measure of reliability that has a value ranging from zero to one (Taber, 2018). Test the validity of each item can be stated as reliable or not based on the comparison between r_{count} and r_{table} value. Items are said to be reliable if the value of $r_{count} > r_{table}$. Based on Guilford's reliability coefficient criteria, which ranges from 0.80 to 1.00, Cronbach's Alpha value must be greater than 0.7 for an instrument to be considered reliable (Prajoko et al., 2021). The results of the Cronbach alpha test with IBM SPSS can be seen in Table 5.

Tabel 5. Reliability Test Results

| Reliability Statistics | |
|------------------------|------------|
| Cronbach's Alpha | N of Items |
| .922 | 5 |

Based on the results of reliability calculations using the alpha cronbach equation, the r_{count} reliability coefficient is 0.922. The r table value used for $n = 18$ with a significance level of 5% is 0.468. With these results, the assessment instrument can be declared reliable because it fulfills $r_{count} > r_{table}$. Assessment instruments that are declared reliable indicate that the questions developed will have the same measurement results even

though they are done at different times, so that these instruments can be used to measure students' argumentation skills repeatedly. These results are in accordance with Putri (2020) research on the development of argumentation skill test instruments which states that five test questions are declared valid and the reliability rate is high at 0.78. With these results, test instruments were developed can be used to measure the argumentation skills on impulse and momentum material.

Conclusion

The theoretical validity of the assessment instrument was assessed based on the content and construct aspects. On content validity, the mode is obtained with a score of 4 and 5 with valid and perfectly valid criteria. In the construct aspect, it gets a mode with a score of 5 with perfectly valid criteria. The empirical validity of the five questions sequentially is 0.791; 0.876; 0.961; 0.927; and 0.871. The r_{count} value meets the $r_{count} > r_{table}$ criteria so that it can be declared empirically valid. Reliability obtained r_{count} value of 0.922. The r_{table} value used for $n = 18$ with a significance level of 5% is 0.468. With these results, the assessment instrument can be declared reliable because it fulfills $r_{count} > r_{table}$. So, the assessment instrument is feasible to measure students' argumentation skills.

Author Contribution

Fitria Rahmatul Ula: Conceptualization, methodology, data curation, writing—original draft preparation, and editing; Suyono: Conceptualization, methodology, formal analysis, validation, and writing—review.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

References

- Abduh, N. K., Sastromiharjo, A., & Anshori, D. S. (2019). Pola Argumentasi Pada Genre Teks Eksposisi Karangan Siswa Sma. *RETORIKA: Jurnal Bahasa, Sastra, Dan Pengajarannya*, 12(1), 71. <https://doi.org/10.26858/retorika.v12i1.7372>
- Afni, A. N., & Suyono, S. (2021). Kelayakan Lembar Penugasan Terstruktur pada Materi Laju Reaksi untuk Melatihkan Literasi Sains. *PENDIPA Journal of Science Education*, 6(1), 16–25. <https://doi.org/10.33369/pendipa.6.1.16-25>
- Ain, T. N., Wibowo, H. A. C., Rohman, A., & Deta, U. A. (2018). The scientific argumentation profile of

- physics teacher candidate in Surabaya. *Journal of Physics: Conference Series*, 997(1). <https://doi.org/10.1088/1742-6596/997/1/012025>
- Amalia, W. N., Amir, H., & Sumpono, S. (2019). Studi Perbandingan Kemampuan Berpikir Kritis Siswa Menggunakan Model Kooperatif Tipe Take and Give Dan Think Pair Square Di Kelas Xi Ipa Sma Negeri 1 Rejang Lebong. *Alotrop*, 3(1), 8–14. <https://doi.org/10.33369/atp.v3i1.9026>
- Ambiyar, & Panyahuti. (2020). *Asesmen Pembelajaran Berbasis Komputer dan Android*. Prenada Media.
- Azwar, S. (2018). *Reliabilitas dan Validitas* (4th ed.). Pustaka Pelajar.
- Dachliyani, L. (2020). Instrumen Yang Sahih: Sebagai Alat Ukur Keberhasilan Suatu Evaluasi Program Diklat (Evaluasi Pembelajaran). *Media Informasi Dan Komunikasi Diklat Kepustakawanan*, 5(1), 57–65. Retrieved from <https://ejournal.perpusnas.go.id/md/article/view/721>
- Devi, N. D. C., VH, E. S., & Indriyanti, N. Y. (2018). Analisis Kemampuan Argumentasi Siswa SMA pada Materi Larutan Penyangga. *JKPK (Jurnal Kimia Dan Pendidikan Kimia)*, 3(3), 152–159. Retrieved from <https://jurnal.uns.ac.id/jkpk/article/viewFile/23308/18616>
- Du, Y., Arkesteijn, M. H., den Heijer, A. C., & Song, K. (2020). Sustainable assessment tools for higher education institutions: Guidelines for developing a tool for China. *Sustainability (Switzerland)*, 12(16). <https://doi.org/10.3390/su12166501>
- Dwiretno, G., & Setyarsih, W. (2018). Pembelajaran Fisika Menggunakan Model Argument Driven Inquiry (Adi) Untuk Melatihkan Kemampuan Argumentasi Ilmiah Peserta Didik. *Inovasi Pendidikan Fisika (IPF)*, 7(2), 337–340. Retrieved from <https://ejournal.unesa.ac.id/index.php/inovasi-pendidikan-fisika/article/view/24505>
- Elyani, E., Sumpono, S., & Amir, H. (2019). Perbandingan Pembelajaran Kooperatif Tipe Structured Numbered Heads (Snh) Dan Numbered Head Together (Nht) Terhadap Kemampuan Berpikir Kreatif Siswa Kelas X Mipa Sman 6 Kota Bengkulu. *Alotrop*, 3(1), 82–90. <https://doi.org/10.33369/atp.v3i1.9046>
- Erlinawati, E., & Muslimah, M. (2021). Test Validity and Reliability in Learning Evaluation. *Bulletin of Community Engagement*, 1(1), 26. <https://doi.org/10.51278/bce.v1i1.96>
- Fidia, F., Puspitawati, R. P., & Yakub, P. (2022). Pengembangan Instrumen Soal Higher Order Thinking Skills (HOTS) Materi Jaringan dan Organ pada Tumbuhan Kelas XI SMA. *Berkala Ilmiah Pendidikan Biologi (BioEdu)*, 11(3), 745–754. <https://doi.org/10.26740/bioedu.v11n3.p745-754>
- Hidayati, R., Kusmanto, A. S., & Kiswanto, A. (2023). Development and Construct Validation of Indonesian Students Self-Confidence Scale Using Pearson Product Moment. *Pegem Journal of Education and Instruction*, 13(3), 94–103. <https://doi.org/10.47750/pegegog.13.03.11>
- Imania, K. A., & Bariah, S. K. (2019). Rancangan Pengembangan Instrumen Penilaian Pembelajaran Berbasis Daring. *Jurnal Petik*, 5(1), 31–47. <https://doi.org/10.31980/jpetik.v5i1.445>
- Irvan, A., & Admoko, S. (2020). Analisis kemampuan argumentasi ilmiah siswa berbasis pola toulmin's argument pattern (tap) menggunakan model argument driven inquiry dan diskusi pada pembelajaran fisika SMA. *Inovasi Pendidikan Fisika*, 9(3), 318–324. Retrieved from <https://ejournal.unesa.ac.id/index.php/inovasi-pendidikan-fisika/article/view/35092/31224>
- Jumadi, J., Perdana, R., Riwayani, & Rosana, D. (2021). The impact of problem-based learning with argument mapping and online laboratory on scientific argumentation skill. *International Journal of Evaluation and Research in Education*, 10(1), 16–23. <https://doi.org/10.11591/ijere.v10i1.20593>
- Kamilahrohrawati. (2018). Validity And Effectiveness Of Student Worksheet Based Toulmin' S Argument. *BIOEDUSAINS: Jurnal Pendidikan Biologi Dan Sains*, 7(3), 450–458.
- Marselina, K. T., Lasmawan, I. W., & Dantes, N. (2021). Pengembangan Instrumen Kemampuan Berpikir Kritis Dan Hasil Belajar Ips Pada Siswa Kelas V Sd. *Jurnal Penelitian Dan Evaluasi Pendidikan Indonesia*, 11(2), 105–114. <https://doi.org/10.23887/jpepi.v11i2.620>
- Permata Sari, I. (2018). Analisis Keterampilan Argumentasi Ilmiah Siswa Kelas Xi Ipa Menggunakan Model Toulmin's Argument Pattern (Tap) Dengan Penerapan Metode Problem Solving. Institut Agama Islam Negeri (Iain) Batusangkar.
- Prajoko, S., Anjani, A., Oktaviani, R. D., Fathimah, P., & Kamaludin, W. (2021). Development of Self-assessment Instruments to Measure Student Affective Domains on Online Biology Learning. *Thabiea : Journal of Natural Science Teaching*, 4(2), 185. <https://doi.org/10.21043/thabiea.v4i2.11964>
- Putri, F. D. (2020). Pengembangan Instrumen Tes Kemampuan Berargumentasi Pada Materi Impuls Dan Momentum. *Khazanah Pendidikan*, 14(1), 44–57. <https://doi.org/10.30595/jkp.v14i1.8463>
- Rafiqoh, S. (2020). Arah Kecenderungan dan Isu Dalam Pembelajaran Matematika Sesuai Pembelajaran Abad 21 Untuk Menghadapi revolusi industri 4.0. *Jurnal MathEducation Nusantara*, 3(1), 58–73.

- Rahayu, Y., Suhendar, & Ratnasari, J. (2020). Keterampilan Argumentasi Siswa Pada Materi Sistem Gerak SMA Negeri Kabupaten Sukabumi-Indonesia (Student's Argumentation Skills on Motion Systems Material at SMA Negeri Sukabumi-Indonesia). *Biodik: Jurnal Ilmiah Pendidikan Biologi*, 6, 312–318. Retrieved from <https://online-journal.unja.ac.id/biodik/article/view/9802>
- Rahmawati, D., & Suprpto, N. (2019). Pengaruh Pembelajaran Guided Discovery Terhadap Keterampilan Argumentasi Tertulis Peserta Didik SMA. *Inovasi Pendidikan Fisika*, 8(3), 891-894. Retrieved from <https://ejournal.unesa.ac.id/index.php/inovasi-pendidikan-fisika/article/view/30598>
- Rizky Hilaldy. (2021). Content Validity of English Summative Test. *Journal of Language Learning and Research (JOLLAR)*, 4(1), 37–41. <https://doi.org/10.22236/jollar.v4i1.7708>
- Sahidu, H., Gunawan, Suranti, N. M. Y., & Nisrina, N. (2020). *Model E-Assessment Dan Implikasinya Dalam Pembelajaran*. CV. Literasi Nusantara Abadi.
- Setiabudi, A., Mulyadi, M., & Puspita, H. (2019). An Analysis of Validity and Reliability of a Teacher-Made Test. *Journal of English Education and Teaching*, 3(4), 522–532. <https://doi.org/10.33369/jeet.3.4.522-532>
- Slamet, R., & Wahyuningsih, S. (2022). Validitas Dan Reliabilitas Terhadap Instrumen Kepuasan Ker. *Aliansi : Jurnal Manajemen Dan Bisnis*, 17(2), 51–58. <https://doi.org/10.46975/aliansi.v17i2.428>
- Syerliana, L., Muslim, & Setiawan, W. (2018). Argumentation skill profile using “toulmin Argumentation Pattern” analysis of high school student at Subang on topic hydrostatic pressure. *Journal of Physics: Conference Series*, 1013(1). <https://doi.org/10.1088/1742-6596/1013/1/012031>
- Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Yusup, F. (2018). Uji Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif. *Jurnal Tarbiyah : Jurnal Ilmiah Kependidikan*, 7(1). <https://doi.org/10.18592/tarbiyah.v7i1.2100>