JPPIPA 9(12) (2023)

Jurnal Penelitian Pendidikan IPA

Journal of Research in Science Education

http://jppipa.unram.ac.id/index.php/jppipa/index

Development of Argumentation Skills Test Instruments on Buffer Solutions in Daily Life

Lilla Pangestu Harwyandani[1], Suyono[2*]

[1] Lilla Pangestu Harwyandani, Chemistry Education Study Program, State University of Surabaya, Surabaya, Indonesia.
[2] Suyono, Chemistry Department, State University of Surabaya, Surabaya, Indonesia.

Received: July 18, 2023
Revised: September 4, 2023
Accepted: December 25, 2023
Published: December 31, 2023

Corresponding Author:
Suyono
suyono@unesa.ac.id

DOI: 10.29303/jppipa.v9i12.4725

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

Abstract: The purpose of this research is to develop a valid and reliable instrument for testing argumentation skills in buffer solutions in daily life. Argumentation skill test instruments are reviewed from theoretical validity (content and construct), empirical validity, and reliability. Development of argumentation skill test instruments using the ADDIE development model with stages: (1) analysis; (2) design; (3) development; (4) implementation; and (5) evaluation. This research was carried out until the initial implementation and then an evaluation was carried out. The argumentation skill test instrument that has been designed is then tested for theoretical validity through expert judgment involving three validators. The argumentation skill test instrument has criteria of theoretical validity if each statement has a mode ≥ 4 with valid categories. After being declared theoretically valid, argumentation skill test instruments are implemented to learners. The results of the argumentation skill test instrument are then analyzed to determine empirical validity and reliability. The results showed that the argumentation skill test instrument was declared theoretically valid where each statement had a mode of 5 with a very valid category. The test instrument is declared empirically valid where each question item has a consecutive calculated r value of 0.807; 0.879; 0.911; 0.911 is declared valid with $r_{count} > r_{table}$ (0.468). The reliability of the argumentation skill test instrument of 0.876 is declared reliable with a $r_{count} > r_{table}$ (0.468).

Keywords: Argumentation Skills Test; Buffer Solution Instrument; Reliability; Validity.

Introduction

A good learning system is a learning system that refers to 21st century learning, namely the ability to think critically, solve problems, collaborate, and communicate. The main component in the 21st century learning process is the involvement of important aspects such as formulating questions, describing mechanisms, and building arguments (Pritasari et al., 2016). As an effort to improve the quality of education in accordance with existing demands, the government has done various ways by issuing guidelines for skills that students must have, according H. Mercier and D. Sperber including 4C competencies (the ability to think critically, think creatively, cooperate with others, and communicate well) (Devi et al., 2018). One of the 4C competencies is critical thinking skills and communication skills. This ability to think critically and communicate skills become a unity in argumentation skills (Devi et al., 2018).

In Umah et al. (2016) Stephen E. Toulmin is a philosopher who advocates an approach to analyze arguments that are different from formal logic approaches through his work in 1958 "The Uses of Argument." According to Toulmin's Argumentation Pattern (TAP) that scientific argumentation is a dialogue between two or more individuals who coordinate facts and theories to provide a more in-depth explanation (Erduran & Osborne, 2004). Argumentation is an attempt to convince someone of the opinion or statement expressed accompanied by factual data (Mc Neill as quoted in Afandi & Rusmini, 2021). Argumentation can be interpreted as proving an argument or opinion that is supported by data, evidence, explanations, or other

How to Cite:
Harwyandani, L. P., & Suyono. (2023). Development of Argumentation Skills Test Instruments on Buffer Solutions in Daily Life. *Jurnal Penelitian Pendidikan IPA*, *9*(12), 10705–10714. https://doi.org/10.29303/jppipa.v9i12.4725

appropriate reference sources (Putri et al., 2020). Argumentation is the process of strengthening a claim through critical thinking analysis based on the support of evidence and logical reasoning. These evidences can contain objective facts or conditions that can be accepted as truth (Ginanjar et al., 2015).

There are several reasons why scientific argumentation skills are very important for students. Such as understanding of concepts, learning quality, and reasoning skills will increase. Argumentation skills can improve critical and logical thinking skills and improve problem-solving skills. Students can build sociocultural activities by interpreting, supporting, or refuting an argument. Students are more confident in conveying ideas because they are based on supporting evidence and explanations (Muna & Rusmini, 2021).

Chemistry is one of the sciences that is considered complex and abstract because it has a combination of material involving three concepts of representation, macroscopic, sub-microscopic, and symbolic (Johnstone, 2009). Chemistry requires high-level cognitive abilities, one of which is buffer solution material. Buffer solution is a material that is often found in daily life including in the body that requires understanding three levels of chemical representation, starting with macroscopic, understanding existing phenomena, submicroscopic and symbolic, namely connecting with the concept of buffer solution. This makes most students lack a thorough understanding of the buffer solution material (Devi et al., 2018). According to research conducted by Marsita et al. (2010) explained that in buffer solution material, the concept that is considered difficult is in the sub-material of the role of buffer solution in the body of living things and daily life.

The ability to argue in the buffer solution is something that needs to be studied further. Argumentation skills can enrich KD 3.12 which contains explaining the working principle, pH calculation, and the role of buffer solutions in the body of living things. Explaining in KD 3.12 means explaining or providing in-depth explanations that are in line with the indicator of argumentation skills, namely warrant. As explained earlier, buffer solution material is material that departs from phenomena and events in life that are close to students.

Buffer solution material requires caution, especially in the process when students connect facts that are often found in their lives with concepts learned in learning. In reality many students are afraid to argue in a learning discussion. Teachers usually only ask students to choose answers, make brief explanations, calculate using a formula, and give conclusions to a material (Afandi & Rusmini, 2021). Students are actually capable in terms of the material taught but most still feel afraid to reveal what is known. Students are also still relatively unaccustomed and not even confident to just present their arguments, even though argumentation skills are needed since the elementary school level so that students are accustomed to expressing opinions or arguing in a discussion forum. In the research of Devi et al. (2018) it is stated that the argumentation ability of students on buffer solution material is still at a low-medium level, which only shows claims or claims with data evidence.

The ability to argue is one of the skills that can be used to hone 21st century skills. The readiness of students who have 21st century skills can be pursued through the education level (Devi et al., 2018). One important component in the education system that can be used to measure the success and achievement of a learning process with evaluation activities. Evaluation activities can be a benchmark for the development of education quality, so as to show the process of student progress towards the goals and provisions set in the curriculum. The results of evaluation activities can be in the form of measurements and assessments that show the success rate of an educational program that has been implemented (Permendikbud, 2017). Through measuring students' argumentation skills on buffer solution material, it can be seen using instruments or measuring instruments. In general, instruments used by teachers in measuring learning outcomes using test instruments. Through test instruments, it is expected to be able to determine the level of ability or potential possessed by students.

Argumentation skills tests on buffer solutions in daily life need to be developed to determine feasibility in terms of validity and reliability. The main characteristics possessed by test instruments are classified into validity, reliability, and level of usefulness (Arifin, 2017). According to Azwar (2011) test instruments have good criteria as measuring instruments are when they are valid and reliable. This test instrument can be used as an alternative assessment with a special purpose to check whether there has been assimilation and accommodation of argumentation skills in the cognitive structure of students.

Method

The type of research used is development research using the ADDIE model developed by Branch (2009) which has five stages. The stages of the ADDIE model are analysis, design, development, implementation, and evaluation. The implementation stage used pilot implementation and the evaluation was a formative evaluation. Analysis and evaluation, A needs analysis was carried out by looking for information about the argumentation skills possessed by students, and

curriculum analysis by determining the KI and KD referring to the 2013 curriculum. The evaluation of this stage is reviewed by the supervisor and revised in accordance with the advice of the supervisor.

Design and evaluation, design stage is carried out by designing test instruments in accordance with information from the analysis stage. At this stage, the design of research instruments is also carried out, namely validity sheets. The evaluation of the design stage was carried out by the supervisor and then revised to improve the design.

Development and evaluation, designed test instruments were later developed. The argumentation skill test instrument contains 4 phenomena of the role of buffer solutions with 6 indicators of argumentation skills as questions on each phenomenon. The validity sheet that has been designed is also developed by containing aspects that will be validated by validators. The evaluation of the development stage is carried out instrument validation by validators who then revise test instruments based on content and construct validity. Implementation and evaluation, Implementation of argumentation skills test instruments for class XI Science in High School as many as 18 students. In evaluation, validity analysis and reliability are carried out based on the results of student answers.

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 analysis technique for content and construct validity is carried out by quantitative descriptive. 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.

Data from expert are then analysis using frequently appearing assessment scores (Mode) obtained from three validators with a minimum mode of 4 (valid) (Afni & Suyono, 2021). Empirical validity of argumentation skill test instruments using the Pearson Product Moment correlation in SPSS 25 is said to be valid if the value of $r_{count} > r_{table}$ (Arikunto, 2005). The reliability of the argumentation skill test instrument uses Alpha Cronbach with reliable criteria if $r_{counts} > r_{table}$ (Arikunto, 2005).

Table 1. Likert scale score

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

(Riduwan & Sunarto, 2017)

Result and Discussion

In education, instruments that used to collect data can be in the form of test or non-test (Jihad & Haris, 2012). The test is a systematic procedure made in the form of tasks that will be given to individuals or groups to be done and responded to both in written, explain, and activity (Matondang, 2009). Test instruments as measuring instruments are used to measure the level of development or progress that has been achieved by students after completing the learning process (Sudjono, 2003).

The argumentation skills test which developed can be used to find out the argumentation skills possessed by each learner. According to Azwar (2011), a good test instrument in this case the argumentation skill test instrument developed must meet the criteria of validity (theoretical and empirical) and reliability as a measuring instrument. The argumentation skill test instrument was developed using the ADDIE development model which has 5 stages, namely: (1) analysis stage, (2) design stage, (3) development stage, (4) implementation stage and (5) evaluation stage (Branch, 2009).

Analysis Result

The analysis stage is carried out to analyze the needs in the learning process so that the argumentation skills test instrument can be applied in schools. Needs analysis is carried out by observation and interviews at school as a means to find information about students' argumentation skills. Based on an interview with a class XI teacher of School that class XI has never applied argumentation skills to the learning process and assessment process. So far, chemistry teachers in schools only use ordinary daily learning and assessments to find out the concepts that students have. In this case, an argumentation skill test instrument was developed on a buffer solution in daily life.

Curriculum analysis is carried out to determine basic competencies and learning objectives so that the argumentation skills test instruments developed are in accordance with the applicable curriculum. The school in class XI uses the 2013 curriculum. Based on the results of an interview with a chemistry teacher that class XI uses the 2013 curriculum, the preparation of

argumentation skill test instruments is adjusted to class XI material in the 2013 curriculum.

Based on curriculum analysis, it was found that the material used in the development of test instruments was buffer solution material in daily life. Argumentation skills on chemistry are implicit in the buffer solution KD 3.12. KD 3.12 explains the working principle, pH calculation, and the role of buffer solutions in the body of living things. In KD 3.12, there is a word explain which can be interpreted to explain or provide a more in-depth explanation that is relevant to the indicator of argumentation skills, namely warrant or giving explanation.

Evaluation at the analysis stage is carried out regarding the use of argumentation skill indicators in the test instrument. The results of the evaluation with the supervisor there was a change for the argumentation skill indicator which originally had 5 indicators, to 6 indicators. This is because the rebuttal indicator is an indicator that can strengthen the arguments of students.

Design Result

The design stage is a stage in the form of making or preparing a design for the development of argumentation skill test instruments.

Test instrument design

At this stage, an initial design of the test instrument to be developed is carried out. The designed test question grid refers to the buffer solution material in daily life. The test instrument grid contains Basic Competencies, namely KD 3.12, aspects of argumentation ability, indicators of argumentation skills, question items, and answers. Below is one of the displays of the grid of argumentation skill test instruments on buffer solutions in daily life. The next stage of planning is to compile the problem. The form of the test to be designed is an essay test with six indicators of argumentation skills. There are four questions developed and each has six indicators of argumentation skills. Indicators of argumentation skills include claims, data, warrants, backing, qualifiers, and rebuttals that contain a refutation or denial of a claim by including reasons. The developed questions contain phenomena regarding buffer solutions in life and include indicators of argumentation skills in each phenomenon. Assessment rubrics can be used to point appropriate learners' answers. This assessment rubric is used to determine scores on each indicator of argumentation skills. Each indicator of argumentation skills has a score of at most 3 and a lowest of 0.

Research instrument design

The design stage of the research instrument used in the development of argumentation skill test instruments is the product validation sheet. The theoretical validity of the argumentation skill test instrument can be seen using a validation sheet which will later be assessed by validators. The theoretical validity of the test instrument will be validated by three validators, two validators come from chemistry lecturers and one validator is a chemistry teacher. The validator will assess the test instrument using a validation sheet in terms of content and construct.

Argumentation skill test instruments that are still in the form of designs or drafts need to be evaluated so that the test instruments developed become better. There are several evaluations from the supervisor that should be improved. After being repaired and having received approval from the supervisor, the test instrument that has been designed can be developed.

Development Result

The development stage is the realization of the draft concept regarding the argumentation skill test instrument that was previously made. The designed test instruments are then developed into instruments that are suitable for use.

Development of test instruments

Instrument cover of argumentation skills

On the outer cover of the test instrument contains the title of the instrument, the name of the compiler, and a place to write the identity of the student.

Figure 1. Cover Display on Argumentation Skills Test Instrument

Instrument description

On the argumentation skill test instrument there is a description of the instrument. In the description of the instrument, students are expected to be able to know and understand the instrument of the argumentation skills test thoroughly. In this section of the description there

are basic competencies that become the material to be achieved. There is also an explanation of the indicators of argumentation skills used in the test instrument and the language structure of each indicator.

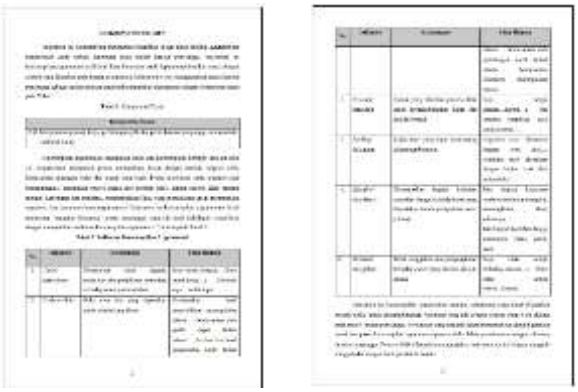

Figure 2. Description of Argumentation Skills Test Instrument

Test questions

In the argumentation skill test instrument, there are four phenomena, each phenomenon has six questions about the indicators of argumentation skills. The argumentation skill test instrument is developed in essay because it can require students to answer in the form of explaining, discussing, comparing, and giving reasons according to the questions given using their own sentences (Sudjana, 2009). Each phenomenon has two claims that will later be chosen by students as appropriate claims. Each indicator also includes instructions to answer so that it is expected to make it easier for students to answer each indicator. Below is a display of phenomena 1 to 4 contained in the designed test instrument.

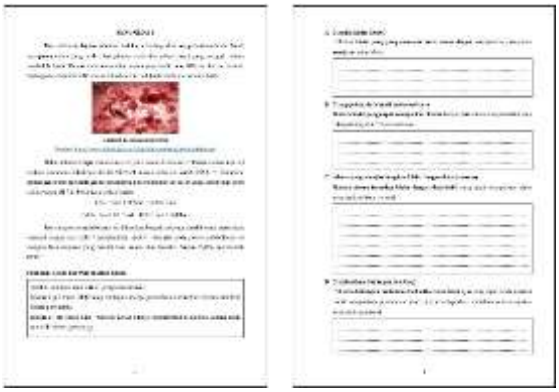

Figure 3. Display of Phenomenon 1 on the Argumentation Skill Test Instrument

Figure 3 shows phenomenon 1 which contains a review of the role of buffer solutions in the blood accompanied by claims that will later be chosen by students.

Figure 4. Display of Phenomenon 2 on the Argumentation Skill Test Instrument

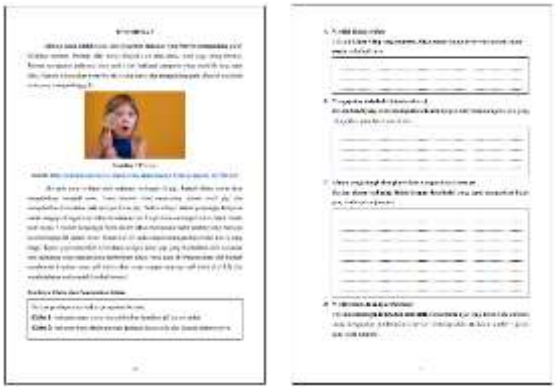

Figure 5. Display of Phenomenon 3 on the Argumentation Skills Test Instrument

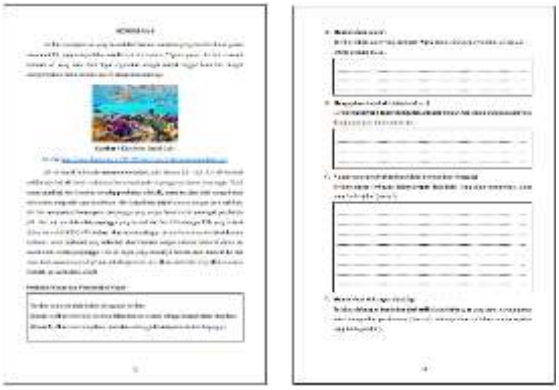

Figure 6. Display of Phenomenon 4 on the Argumentation Skills Test Instrument

Figure 4 shows phenomenon 2 which contains a review of the role of buffer solutions in the blood accompanied by claims that will later be chosen by students. Figure 5 shows phenomenon 3 which contains a review of the role of buffer solutions in saliva accompanied by claims that will later be chosen by students. Figure 6 shows phenomenon 4 which contains a review of the role of buffer solutions in water accompanied by claims that will later be chosen by students.

Assessment rubrics

This scoring rubric is used to determine scores on each indicator of argumentation skills. Below is the scoring rubric for the argumentation skills test instrument prepared.

Figure 7. Argumentation Skills Test Instrument Assessment Rubric

Research instrument development

The test instrument validation sheet consists of content validation and constructs regarding the argumentation skill test instrument developed. Validity is a term that refers to the concept of whether a test can measure something that has been planned to be measured (Suryabrata, 2005). Assessment instruments have a level of feasibility that can be seen from theoretical validity and empirical validity (Arikunto, 2015). The validity of the argumentation skill test instrument is reviewed from theoretical and empirical validity. Argumentation skill test instruments before implementation must first be validated by validators.

Theoretical validity is a condition where the instrument developed meets valid requirements based on existing theories and provisions (Newman et al., 2013). Theoretical validity consists of content and construct validity. Content validity refers to an instrument developed in accordance with predetermined criteria or the extent to which the instrument can include content that must be measured (Arnilawati et al., 2018). Construct validity means the extent to which a measuring instrument can measure the construct of something being measured (Anastasi & Urbina, 1997). Construct validity is a measurement that correlates one indicator with another indicator and is associated with variables contained in the construct theory to be measured (Westen & Rosenthal, 2003). The validity of the construct can also be known by using how to detail and pair each question item with every aspect in the indicator to be used through the aspect of the framework arrangement (Nizary & Kholik, 2021). The argumentation skill test instrument is considered valid if each statement has an assessment mode with a minimum score of 4 which has valid criteria (Afni & Suyono, 2021).

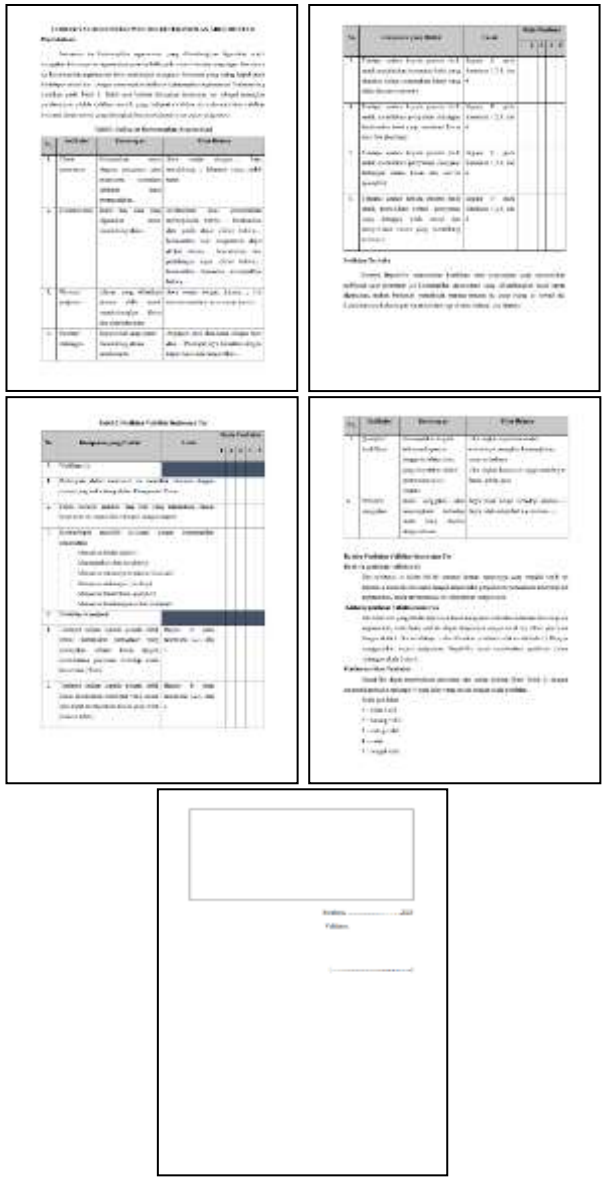

Figure 8. Argumentation Skills Test Instrument Validation Sheet

The argumentation skill test instrument that had been prepared was then validated by three validators consisting of two validators, namely UNESA Chemistry Education lecturers (validators 1 and 2) and one high school Chemistry teacher validator (validator 3). Validation of argumentation skill test instruments in this stage is theoretical validation consisting of validation of the content and construct of the test instrument developed. Each aspect has a maximum value of 5 and a minimum of 1. Validators can also provide advice and input regarding the test instruments prepared. Suggestions and comment from validators can be used

to improve the argumentation skills test instrument to make it better.

Table 2. Instrument Validation Results of Argumentation Skills Test

Assessed Components	Validator 1	2	3	Mo	Validity Criteria
Content validity					
1	4	5	5	5	Very valid
2	4	5	5	5	Very valid
3	5	5	5	5	Very valid
Construct validity					
4	5	5	5	5	Very valid
5	5	5	5	5	Very valid
6	4	5	5	5	Very valid
7	5	5	5	5	Very valid
8	4	5	5	5	Very valid
9	5	5	5	5	Very valid

Based on Table 2 which show the result of theoretical validity from validator, that is can show with graph.

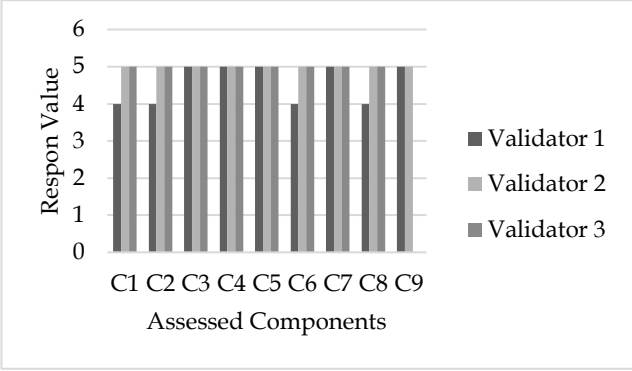

Figure 9. Instrument Validation Results of Argumentation Skills Test

Based on Table 2 and Graph 1, it can be seen that the assessment of the validity results on each statement regarding the validity of content and construct has mode 5 with very valid criteria. The test instrument can be declared valid if each statement has at least 4 modes with valid criteria. It can be concluded that the argumentation skill test instrument on the buffer solution in daily life is declared theoretically valid with very valid criteria.

The argumentation skill test instrument has theoretically valid criteria based on the validity of the content reviewed from the validator's assessment through expert judgment to get a score mode of 5 with very valid criteria. The test instruments developed are valid in accordance with the material contained in the Basic Competencies and have relevance to argumentation skills. This explains that the test instruments developed refer to the material that has been determined in the Basic Competencies. According to Hamzah & Muhlisrarini (2016) there is a way to write tests well, namely by determining the objectives of the learning you want to measure. Facts, concepts, images, and theories contained in the test instrument have relevance to the material as stated by Kastina (2017) that the test instrument used must be in accordance with the characteristics of the competence to be achieved.

Table 3. Assessed component

Numb.	Assessed Components Content validity
1	Questions in the test instrument have relevance to the material contained in the Basic Competencies
2	The facts, concepts, images, and theories contained in the test instrument have relevance to the material.
3	Content/topic has relevance to argument skills 1. Drafting a claim 2. Showing data (evidence) 3. Drafting reasons/justifications (warrant) 4. Backing 5. Develop qualifications (qualifiers) 6. Constructing counterargument (rebuttal)
Numb.	Construct Validity
4	There is a direction for students to write a statement that is a claim by assessing a phenomenon (claim)
5	There are directions to students to write appropriate evidence/data and can strengthen the claims that have been compiled (data)
6	There is a direction to students to explain the suitability of the evidence submitted in strengthening the claims that have been prepared (warrant)
7	There are directions to students to write statements of support based on the theory underlying the claim or data (backing)
8	There are instructions for students to write statements about the relationship between claims and warrants (qualifier)
9	There is a direction for students to write a statement that is considered inappropriate and include supporting reasons (rebuttal)

The argumentation skill test instrument has theoretically valid criteria based on construct validity in terms of validator assessment through expert judgment which gets a score mode of 5 with very valid criteria. The test instruments developed are valid according to construction in accordance with the characteristics of the competencies to be achieved, in this case the indicators of argumentation skills (Kastina, 2017). The test instrument developed contains indicators of argumentation skills and directions that can make it easier for students to answer, so that the argumentation skills test instrument can provide direction to students

to construct knowledge about argumentation skills in buffer solutions in everyday life.

The assessment conducted by validators shows good results, but there are several aspects that need to be improved in order to produce valid test instruments. Comments, suggestions and input from validators can be used to revise argumentation skill test instruments so that theoretically valid argumentation skill test instruments can be produced.

Implementation Result

The implementation of the argumentation skill test instrument was applied to 18 students of grade XI Science SMA in Jombang. The implementation of this argumentation skill test instrument will be carried out on May 31, 2023, which will be carried out for 2 x 40 minutes. Students work on an argumentation skill test instrument containing 4 phenomena with 6 questions which are indicators of argumentation skills (claim, evidence, warrant, backing, qualifier, and rebuttal). Before starting to take the test, students are given directions or instructions regarding filling out test instruments. Students are also given direction to do the questions contained in the argumentation skill test instrument. Students are then given time to work on the argumentation skills test instrument independently.

At the implementation evaluation stage, student answers are produced and then assessed based on the assessment rubric so that each student's score and score are produced. The results of the argumentation skill test instrument that have been done by students that most students are able to complete the test instrument and get a good score. The average score for the test instrument obtained by students is 88.89.

An instrument can be said to have empirical validity if it has been tested from experience. Empirical validity cannot be obtained simply by compiling instruments based on such logical validity, but must be proven through experience (Newman et al., 2013). Empirical validity is obtained through the results of test trials to respondents (Matondang, 2009). Determine the question items that are declared valid based on the implementation results. The empirical validity of test instruments can be seen through the results of the work on the test instruments given to students.

Empirical validity is determined using the Pearson Product Moment correlation coefficient equation. The result of the calculation is compared with the r value in the Product Moment correlation table. The argumentation skill test instrument is said to be empirically valid if the value of $r_{count} > r_{table}$ (Arikunto, 2005). The argumentation skill test instrument was implemented to 18 students.

Table 4. Empirical Validity Test Instrument Result

Phenomenon	r count	Category
F1	0.807	Valid
F2	0.879	Valid
F3	0.911	Valid
F4	0.911	Valid

Based on the explanatory description of the correlation of phenomenon 1, phenomenon 2, phenomenon 3, and phenomenon 4 with a total score, it can be concluded that the test instrument can be said to be empirically valid. Each question item has value of r_{count} for phenomenon 1 of 0.807, phenomenon 2 of 0.879, phenomenon 3 of 0.911, and phenomenon 4 of 0.911. As for the r table with n = 18 and the 5% significance level is 0.468. The four correlated question items fall into the valid criteria with high categories. Four questions that have been developed meeting empirical validity criteria can be declared valid and suitable for use (Arikunto, 2005).

Test reliability shows the level or degree of consistency of an instrument, whether a test is thorough and reliable in accordance with predetermined criteria (Arifin, 2017). An instrument can be said to be reliable or consistent if used many times to measure the same object produced the same data (Sugiyono, 2014). The reliability of the argumentation skill test instrument can be calculated using Cronbach's Alpha equation. Test instruments can be said to have good reliability if the value of $r_{count} > r_{table}$ (Arikunto, 2005).

Table 5 Reliability of Test Instrument

Reliability Statistics	
Cronbach's Alpha	N of Items
.876	4

Based on SPSS calculations regarding the reliability value of argumentation skill test instruments carried out by students, the value of the instrument reliability coefficient or r_{count} at 0.876. As for the r value of the $_{table}$ used for n= 18 with a significance level of 5%, which is 0.468. With the calculation results, it can be stated that the argumentation skill test instrument on the buffer solution in daily life that has been developed can be said to be reliable because the value of $r_{count} > r_{table}$ (Arikunto, 2005).

Conclusion

Based on the results of data analysis and discussion of research on the validity and reliability of argumentation skill test instruments on buffer solutions in daily life, it can be concluded, the validity of the test instrument in terms of the theoretical validity of each statement regarding the validity of the content and construct has mode 5 with very valid criteria and

empirical validity of the test instrument based on the results of students' answers, each question factor has value of r_{count} for phenomenon 1 of 0.807, phenomenon 2 of 0.879, phenomenon 3 of 0.911, and phenomenon 4 of 0.911 where $r_{count} > r_{table}$ which falls into the valid criteria with a high category. The reliability of the argumentation skill test instrument based on the score of the test results done by students obtained the value of the instrument reliability coefficient or r_{count} of 0.879 so that it can be said to be reliable because the value of $r_{count} > r_{table}$.

Author Contribution

Lilla Pangestu Harwyandani: Conceptualization, methodology, data curation, writing original draft preparation, and editing. Suyono: Conceptualization, methodology, formal analysis, validation, and writing review.

Funding

The research received no external funding.

Conflict of Interest

The authors declare no conflict of interest.

References

Afandi, A. A., & Rusmini. (2021). Kelayakan Lembar Kerja Peserta Didik untuk Melatihkan Keterampilan Argumentasi Peserta Didik SMA Kelas XI. In *UNESA Journal of Chemical Education*, *10*(2). https://doi.org/10.26740/ujced.v10n2.p172-184

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

Anastasi, A., & Urbina, S. (1997). *Psychological Testing 7th ed* (7th ed.). Prentice Hall/Pearson Education.

Arifin, Z. (2017). Kriteria Instrumen dalam suatu Penelitian. *Jurnal THEOREMS (The Original Research of Mathematics)*, *2*(1), 28–36. https://doi.org/10.31949/th.v2i1.571

Arikunto, S. (2005). *Dasar-dasar Evaluasi Pendidikan*. Bumi Aksara.

Arikunto, S. (2015). *Dasar-dasar Evaluasi Pendidikan*. Bumi Aksara.

Arnilawati, A., & Musdi, E. (2018). Students' worksheet validity based on contextual teaching and learning in junior high school. In *International Conferences on Educational, Social Sciences and Technology,* 496–500. https://doi.org/10.29210/2018172

Azwar, S. (2011). *Reliabilitas dan Validitas*. Pustaka Belajar.

Branch, R. M. (2009). *Instructional design: The ADDIE approach.* In Springer US. https://doi.org/10.1007/978-0-387-09506-6

Devi, N. D. C., Susanti VH, E., & Indriyanti, N. Y. (2018). Analysis of High School Students' Argumentation Ability in the topic of Buffer Solution. *JKPK (Jurnal Kimia Dan Pendidikan Kimia)*, *3*(3), 152–159. https://doi.org/10.20961/jkpk.v3i3.23308

Erduran, S., & Osborne, J. (2004). TAPping into Argumentation: Developments in the Application of Toulmin's Argument Pattern for Studying Science Discourse. *Science education, 88*(6), 915-933. https://doi.org/10.1002/sce.20012

Ginanjar, W. S., Utari, S., & Muslim, Dr. (2015). Penerapan Model Argument-Driven Inquiry dalam Pembelajaran IPA untuk Meningkatkan Kemampuan Argumentasi Ilmiah Siswa SMP. *Jurnal Pengajaran Matematika Dan Ilmu Pengetahuan Alam*, *20*(1), 32–37. https://doi.org/10.18269/jpmipa.v20i1.559

Hamzah, A., & Muhlisrarini. (2016). *Perencanaan dan Strategi Pembelajaran Matematika* (3rd ed.). Rajagrafindo.

Jihad, A., & Haris, A. (2012). *Evaluasi Pembelajaran*. Multi Pressindo.

Johnstone, A. H. (2009). Multiple Representations in Chemical Education. *International Journal of Science Education*, *31*(16), 2271–2273. https://doi.org/10.1080/09500690903211393

Kastina, Z. V. K. (2017). Implementasi Sistem Penilaian dalam Kurikulum 2013 di SMA Negeri 2 Pekanbaru. In *JOM FISIP*, *4*(1). Retrieved from https://jom.unri.ac.id/index.php/JOMFSIP/article/view/12720

Marsita, R. A., Priatmoko, S., & Kusuma, E. (2010). Analisis Kesulitan Belajar Kimia Siwa SMA dalam Memahami Materi Larutan Penyangga dengan Menggunakan Two-Tier Multiple Choice Diagnostic Instrument. In *Jurnal Inovasi Pendidikan Kimia*, *4*(1). https://doi.org/10.15294/jipk.v4i1.1308

Matondang, Z. (2009). Validitas dan reliabilitas suatu instrumen penelitian. *Jurnal tabularasa, 6*(1), 87-97. Retrieved from http://digilib.unimed.ac.id/id/eprint/705

Muna, A. N., & Rusmini. (2021). Pengembangan Lembar Kerja Peserta Didik untuk Melatihkan Keterampilan Argumentasi Ilmiah Peserta Didik pada Materi Laju Reaksi. In *UNESA Journal of Chemical Education*, *10*(2). https://doi.org/10.26740/ujced.v10n2.p159-171

Newman, I., Lim, J., & Pineda, F. (2013). Content Validity Using a Mixed Methods Approach: Its Application and Development Through the Use of a Table of

Specifications Methodology. *Journal of Mixed Methods Research*, *7*(3), 243–260. https://doi.org/10.1177/1558689813476922

Nizary, M. A., & Kholik, A. N. (2021). Validitas instrumen assesmen (Analisis validitas isi dan konstruk instrumen asesmen buku pelajaran Al Quran hadis kelas 6 madrasah ibtidaiyah materi surat Ad Dhuha Bab Vi). *CONTEMPLATE: Jurnal Ilmiah Studi Keislaman*, *2*(1), 21–42. Retrieved from https://ejournal.iaiqi.ac.id/index.php/contemplate/article/view/49

Permendikbud. (2017). *Penilaian Hasil Belajar Oleh Pemerintah dan Penilaian Hasil Belajar Oleh Satuan Pendidikan*. Direktorat Utama Pembinaan dan Pengembangan Hukum Pemeriksaan Keuangan Negara Badan Pemeriksa Keuangan

Pritasari, A. C., Dwiastuti, S., & Probosari, R. M. (2016). Peningkatan kemampuan argumentasi melalui penerapan model problem based learning pada siswa kelas X MIA 1 SMA Batik 2 Surakarta Tahun Pelajaran 2014/2015. *Jurnal Pendidikan Biologi, 8*(1), 1-7.

Putri, P. A. W., Rahayu, S., & Fajaroh, F. (2020). Efektivitas argument-driven inquiry untuk meningkatkan keterampilan berargumentasi ilmiah pada materi laju reaksi. *Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan, 5*(1), 57-64. Retrieved from http://journal.um.ac.id/index.php/jptpp/article/view/13132

Riduwan, & Sunarto. (2017). *Pengantar Statistika untuk Penelitian: Pendidikan, Sosial, Ekonomi, Komunikasi, dan Bisnis* (8th ed.). Alfabeta.

Sudjana, N. (2009). *Penilaian Hasil Belajar Proses Belajar Mengajar*. PT Remaja Rosdakarya.

Sudjono, A. (2003). *Pengantar Evaluasi Pendidikan*. PT Raja Grafindo Persada.

Suryabrata, S. (2005). *Pengembangan Alat Ukur Psikologis*. Andi.

Umah, U., Asari, A. R., & Sulandra, I. M. (2016). Struktur Argumentasi Penalaran Kovariasional Siswa Kelas Viiib Mtsn 1 Kediri. *JMPM: Jurnal Matematika dan Pendidikan Matematika, 1*(1), 1-12. https://doi.org/10.26594/jmpm.v1i1.498

Westen, D., & Rosenthal, R. (2003). Quantifying Construct Validity: Two Simple Measures. *Journal of Personality and Social Psychology*, *84*(3), 608–618. https://doi.org/10.1037/0022-3514.84.3.608