Journal of Innovative Physics Teaching
Volume 2, Issue 2 December 2024, 89 -101
ISSN: 3021-8551 (Online)
Doi: https://doi.org/10.24036/jipt/vol2-iss2/68
https://jipt.ppj.unp.ac.id/index.php/jipt

Development of Physics Test Instruments to Measure Problem Solving Skills on Dynamic Fluid Materials

Aditya Putra Sanjaya, Yul Ifda Tanjung*, Satria Mihardi
[1] Physics Department, Faculty of Mathematics and Natural Science, Universitas Negeri Medan, Medan, Indonesia.

Article Info

Article history:

Received July 05, 2024
Revised December 13, 2024
Accepted December 20, 2024

Keywords:

Test Instruments
Problem Solving
4D Model
Dynamic Fluids

ABSTRACT

Problem solving skills are a fundamental part of physics learning that must be possessed by students in facing challenges and problems that involve critical thinking, logical, systematic and able to make the right decisions. In fact, the availability of instruments to measure these skills is still limited. Based on that, the purpose of this research is to develop a test instrument based on problem solving skills based on Heller's stages on dynamic fluid material that meets the criteria for good instrument feasibility. This type of research is development research (Research and Development / R&D) with a 4D model consisting of 4 stages, namely define, design, develop and disseminate. The subjects of this study were XI grade high school students totaling 36 people for product testing. The results showed that the average value of content validity testing was 93.53% with very valid criteria, while the results of item validation in small and large groups showed that all ten test instruments were categorized as valid. In terms of reliability, the small group obtained a value of 0.84 which is classified as very high reliability and in the large group obtained a value of 0.78 which is classified as high reliability. The results of the student response test obtained a positive response in the small group of 81.00% with a good category and in the large group of 81.32% with a very good category. Based on the results of the study, it is concluded that the dynamic fluid material problem solving test instrument developed as many as 10 questions has met the eligibility criteria of the test instrument including validity and reliability so that it can then be used as a training medium and a measuring tool for problem solving skills.

***Correspondence:**
Yul Ifda Tanjung, Physics Department, Faculty of Mathematics and Natural Science, Universitas Negeri Medan, Medan, Indonesia.· email : yuly@unimed.ac.id,

INTRODUCTION

Education plays a very important role in every era's development and provides answers to problems that arise along with the times (Suprapto et al., 2023; Ainun et al., 2024). The goals of education include creating a creative, competent and knowledgeable generation (Tanjung et al., 2023). The changes that occur are an effort to improve the quality of education and create a generation of people who have good quality human resources and can compete with other countries. (Tugiah and Jamilus, 2022). Apart from that, education is also expected to create students who have a tough, creative, independent and responsible attitude and are able to blend in with increasingly rapid developments (Sugiarto and Farid, 2023). This is in accordance with the view of Sani (2019) who states that it is very important to equip students to have high creativity, be flexible in critical thinking, be able to make careful decisions and be able to solve existing problems. In line with developments over time, the curriculum has also developed to meet educational demands.

To find out the extent to which educational program organizers are able to achieve their goals carefully and on time, an assessment is therefore needed (Arikunto, 2021). Assessment as a procedure that provides a description of the individual being assessed and considers the values and meaning contained therein (Maison, Astalini & Perdana, 2020; Mohan, 2023; Bozorova & Shaxnoza, 2023). Evaluation is a decision-making activity based on obtaining predetermined measures and criteria (Kopnina, 2020). Measurement and evaluation are two interconnected activities. In learning physics, each student is not only focused on having the ability to understand concepts (low level thinking) but is also expected to have high level thinking abilities, so that they can solve various problems in real life which generally require high level thinking abilities (Mayasari et al., 2023 ; Ratnasari et al., 2021).

Higher Order Thinking Skills (HOTS) include aspects of critical thinking, creativity, decision making and problem solving. Problem solving skills are one of the most important competencies that must be instilled in all students (Garcia, Argelagos & Privado, 2021; Yanto, Festiyed & Enjoni, 2020). Learning based on problem solving skills is designed based on an active, student-oriented learning system, increased curiosity and HOTS-based evaluation (Rapih and Sutaryadi, 2018). Problem solving skills are seen as a fundamental part of physics learning (Jayadi et al., 2020). This is because problem solving skills will stimulate students to think about solving problems using appropriate principles and laws. The importance of having problem solving abilities in students will add new experiences in finding solutions to existing problems.

The physics problem solving process effectively guides students to be able to identify, make decisions and provide problem solving paths using logical thinking, literacy and high creativity (Purwaningsih et al., 2020; Sari, Jatmiko & Suprapto, 2023). Students can be declared capable of problem solving if students are able to determine the basis of the problem to be solved, after that they must be able to determine effective and efficient stages so that the solution can be implemented immediately (Rudnick, 1980; Ridho et al., 2020). Skills in problem solving are solving complex problems where the thinker must be able to analyze and synthesize which is a high-level skill (HOTS) (Winarti et al., 2019; Hobri et al., 2020). When the problem given to the student is something that is well known, then the individual can solve the problem without using problem solving skills and the problem is not a "problem" for the student.

The importance of applying problem solving in the learning process, because in the problem-solving process students are not only required to listen, write and remember learning concepts, but more than that students will be trained to think actively, communicate, collect and process information, and draw conclusions. The capacity to solve complex problems is considered an important skill for every individual in the current technological era. Physics problems provide a good opportunity for students to participate directly in problem solving (Marlina et al., 2024). Apart from studying physics material and principles, skills in solving physics problems are also the main goal, both at school and in college (Sari, Jatmiko & Suprapto, 2023). Problem solving skills are not inherited skills, but rather are skills that can be developed and trained through learning and assessment or evaluation procedures. High-level thinking skills are needed in learning evaluation by getting students used to working on test instruments based on problem-solving skills, which in itself will improve students' problem-solving skills and will reflect directed and organized thinking skills in solving problems.

In fact, physics lessons given to high school students mostly use question instruments that only test the level of concept understanding skills (Rahmi et al., 2024). This is because not all educators can develop appropriate questions to assess higher-level thinking abilities (Rusdianto, 2020). It is stated that the current characteristics of physics learning are focused on procedural skills, communicating material in one direction, monotonous class layout, low thinking ability, focus on books from school, the instruments used are mostly low level Assessments carried out by educators also usually assess more low-level thinking abilities, namely at level 1 and level 2 only, for example memorizing in completing physics test instruments without understanding the material in depth, therefore students' thinking abilities are difficult to develop (Widana, 2017).

Based on the results of interviews and observations of the test instruments applied at SMA Negeri 15 Medan, it is known that most of the test instruments taken by students tend to focus on textbooks, namely test instruments in the form of essays and multiple choice. The test instruments contained in textbooks still use cognitive test instruments which are classified at cognitive levels C1, C2, C3 and C4, while there are still few question instruments with C5 and C6 abilities and questions with problem solving abilities are still rarely applied. Based on research results, 50% students are in low problem-solving skills and only 8% in high problem-solving skills (Tanjung et al., 2024). The results of the above observations were also confirmed in the results of an interview with one of the physics teachers at SMA Negeri 15 Medan. It was discovered that the teacher had begun to apply several daily test instruments in the form of essays with problem solving in him. However, the problem-solving test instrument provided does not fully implement the correct solution steps in accordance with a clear theory. The solutions carried out are only limited to understanding the problem and mathematical calculations. Based on this, the development of test instruments based on problem solving skills is very necessary. The aim of this research is to develop a test instrument based on problem solving skills in dynamic fluid material that meets the criteria for good instrument feasibility.

METHODS

This research is a type of research and development (R&D) by implementing the 4D development model by Thiagarajan which consists of four stages, including the definition stage, the design stage, the development stage, and the disseminate stage. The research was conducted at SMA Negeri 15 Medan in even semester. The research subjects totaled 36 people, with a sample of 18 people in class XI MIPA 4 and 18 people in class XI MIPA 5. The data collection techniques used were interviews, problem solving skills tests, observation and questionnaire/response questionnaire methods. Data analysis used in this research is qualitative and quantitative data analysis.

Qualitative data analysis was obtained from the results of content validity by the assessment of experts/practitioners in terms of material, construction and language. This expert validation is carried out by looking at the suitability between the content of the instrument material with basic competencies and question indicators, the suitability between the construction of questions with indicators of physics problem solving skills, and the suitability of grammar with good and correct rules.

The data from the experts' assessment will be calculated and analyzed to determine the criteria for the validity level of the test instrument developed. Calculation of validation from validators using the formula, as follows:

$$V_a = \frac{T_{sa}}{T_{sh}} x\ 100\% \quad (1)$$

Description:
V_a = Validation Score
T_{sa} = Total empirical score from experts
T_{sh}= Total maximum score

Expert validity criteria are presented in Table 1.

Table 1. Expert Validity Criteria

No	Percentage (%)	Validity Criteria
1	86-100	Very Valid
2	71-85	Valid
3	50-70	Less Valid
4	<50	Invalid

A good test instrument as a measuring tool if it has met the feasibility of a good test item, which has validity and reliability (Mata-López et al., 2021). Quantitative data analysis is obtained from data from product trial results to see the level of item validity and reliability of test instruments. Item Validity Test is a statistical test used to determine how valid a question item is in measuring the variable being studied. Testing the validity of the items in this research was carried out after the test instrument had been developed and trials would be carried out twice, namely the first trial and the second trial. The validation results can be stated in the category of valid or not after the validation results are obtained in accordance with the specified criteria (Rahmi et al., 2024). To test the validity of this item, the Pearson Product Moment Correlation Test will be used. In this test, each item will be tested for its relationship with the total score of the variable in question. In this case, each item in variables X and Y will be tested for its relationship with the total score for that variable. The validity level of the instrument was obtained using the Product Moment correlation test using the equation:

$$r_{xy} = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{\{N\sum X^2 - (\sum X^2)\}\{N\sum Y^2 - (\sum Y^2)\}}} \quad (2)$$

Item Reliability Test is a statistical test used to determine the reliability of a series of question items in terms of their reliability in measuring a variable. Test reliability regarding the question, whether a test is thorough and reliable in accordance with the established criteria. A test is said to be reliable if it always gives the same results when tested on the same group at different times or on different occasions. Reliability test analysis is carried out to determine the consistency of the test using the Cronbach's Alpha test formula using the equation:

$$r_{11} = \left(\frac{k}{k-1}\right)\left(1 - \frac{\sum s_i^2}{s_i^2}\right) \quad (3)$$

The analysis of learner responses is calculated based on the percentage of positive or negative responses to the test instrument products that are carried out. The data analysis technique used in this research is instrument testing which is tested based on questionnaires or questionnaires that have been distributed to respondents who have worked on the distributed test instruments. The test results of filling out the response questionnaire and the observations that have been obtained will provide an idea of how students respond to the problem-solving skills instrument. The percentage calculation is done using the equation:

$$\%\, response = \frac{score\ obtained}{maximum\ score} x\ 100 \qquad (4)$$

The criteria for analyzing students' responses to the problem-solving skills-based test instrument are presented in Table 2.

Table 2. Student Response Questionnaire Criteria

No	Percentage (%)	Criteria
1	76-100	Very Valid
2	51-75	Valid
3	26-50	Less Valid
4	0 – 50	Invalid

(Sudijono, 2008)

RESULTS AND DISCUSSION

Results

The results of the test instrument development process are obtained based on the success of each stage of the development process that has been determined previously, namely using the 4-D (Four D) model. The results obtained at each stage of the development process are described in the following paragraphs. At the initial stage, namely define, some data or information regarding the potential and existing problems were obtained, including the results of needs analysis, learner analysis, task analysis, and concept analysis. The results of the needs analysis showed that the need for a physics problem-solving skills-based test instrument to measure and see the extent of students' problem-solving skills in solving physics problems at SMA Negeri 15 Medan school is needed. This is clarified in the results of observations and interviews regarding the evaluation and assessment carried out by educators have not fully implemented the assessment based on higher order thinking skills.

The results of the analysis of learners obtained a review of the characteristics of learners including background knowledge and level of cognitive development. The skills possessed by students vary, namely there are high, medium and low abilities. This is due to several underlying factors such as parenting, intelligence or intelligence, motivation, interest or talent and their attitude towards their surroundings. The results of the task analysis show that the tasks given by teachers are in the form of essays and multiple choices, which tend to use questions that only measure students' learning outcomes. Educators are more dominant in giving assignments to students at the cognitive level at the level of understanding to analysis only which is still classified in the Lower Order Thinking and Higher Order Thinking categories. Therefore, there is a need for a shift towards encouraging critical thinking and problem-solving skills in students to foster deeper learning.

The results of the material analysis are carried out by identifying the main concepts that will be applied by analyzing the competency standards and basic competencies of high school physics learning sourced from the syllabus. Based on the material analysis activities on the syllabus, it is found that the material to be used in the development of problem solving skills-based test instruments is dynamic fluid material with several main sub-topics including the continuity equation, Bernoulli's law, Torricelli's theorem and the application of dynamic fluid in the field of technology. The design stage in the development of test instruments includes the preparation of test specifications (including determining basic competencies, problem indicators, indicators of problem-solving skills, determining the form of instruments in the form of descriptions, adjusting to the subject matter and preparing grids); designing question instruments along with answer criteria; making scoring guidelines; and making expert validation sheets. This process ensures that all necessary aspects of the test development are covered, from the conceptual framework to the final validation steps. The goal is to create a comprehensive, reliable, and valid test instrument that can accurately assess problem-solving skills in the context of dynamic fluid material.

The preparation of question indicators is made according to and refers to basic competencies. The question indicators in the development of this test instrument product are as follows: (1) Analyze the equations in the Continuity Principle and Bernoulli's Law to solve problems in everyday life; (2) Evaluate problems related to the design of aircraft wings and mosquito sprayers; (3) Analyze the Torricelli equation related to leaky tank problems in everyday life. The five indicators of physics problem solving skills that must be fulfilled based on Heller's theory (2010), are as follows: (1) Focus the Problem; (2) Describe the Problem in Terms of Physics; (3) Plan a Solution; (4) Execute the Plan; (5) Evaluate the Answer. At the develop stage, the product of the initial design of the test instrument that has been completed will be validated by the validator and then will be tested on the research subject as a limited trial until the product is produced in accordance with the predetermined specifications. Based on the processed expert validation results, the average percentage calculation is 93.53% and after being matched with the validity criteria, it is classified as a very valid category.

Table 3. Results of Expert Validity Analysis of Each Test Item

Question Number	Average Validator Score	Validity Criteria
1	93.63%	Valid
2	93.63%	Valid
3	95.10%	Valid
4	93.14%	Valid
5	92.65%	Valid
6	93.63%	Valid
7	93.14%	Valid
8	91.67%	Valid
9	94.12%	Valid
10	94.61%	Valid

Based on content validation test calculations from the expert panel assessment, it shows that of the 10 questions reviewed by 3 expert validators in the field of assessment, it was found that the ten questions were declared valid. This is proven by the average value of the percentage calculation reaching > 90% and after being matched with the validity criteria it is classified as very valid with the requirement for slight revisions as suggested by the validator. It can be concluded that the test instrument based on problem solving skills on the subject matter of fluids is feasible. Based on the three main aspects that need to be considered in development test instrument, an assessment analysis is carried out to see the level of validity of the test instrument on material, construction and language aspects. The results of the analysis will then be classified to determine whether the test instrument is valid or not in three aspects. Test instruments that are classified as valid can be used for testing with several minor revisions. The results of the assessment based on the three aspects of the test instrument in terms of material, construction and language for each validator are presented in Table 4.

Table 4. Percentage of Validation Assessment of the Three Aspects of the Test Instrument

Aspects	Average Percentage	Validity Criteria
Material	88%	Valid
Construction	85.5%	Valid
Language	85.25%	Valid

The results of the analysis of the three main aspects in developing the test instrument showed that the average percentage for the material aspect was 88% classified as valid, 85.5% for the construction aspect classified as valid and 85.25% for the language aspect classified as valid. Based on this percentage, a total of 10 test items were obtained that met the three aspects, namely material, construction and language with an average of being in the valid category, however revisions or improvements were still carried out based on suggestions and comments from the expert panel (validators). All test items can then be continued at the product testing stage. This reveals questions that are suitable in terms of content validity and can be used as a subsequent test or at a later date. The results of the feasibility of the test instrument items are obtained based on the results of product testing in small and large groups to determine the level of item validity and reliability of the test instrument. The results of test validity in small and large groups can be seen from the distribution table of the 10 items based on the results of item validity analysis, as presented in Table 5.

Table 5. Validity of Test Instrument Items in Small Group

Question Number	r_{count}	r_{table}	Validity Criteria
1	0.649		Valid
2	0.659		Valid
3	0.714		Valid
4	0.682		Valid
5	0.681	0.631	Valid
6	0.690		Valid
7	0.650		Valid
8	0.696		Valid
9	0.649		Valid
10	0.633		Valid

Based on the results of the validity test analysis in small groups with a total of 10 participants and the r-table value showing 0.631 at a significance level of 5%, it was found that 10 question items were declared to be in the valid category. This is appropriate based on the correlation coefficient value r calculated > r table, so the test instrument items are significantly correlated (declared valid). Valid test items reflect that the test instrument has reliability and there is no doubt about the accuracy of the test instrument in measuring students' abilities. The instrument has high validity if the items are forming an instrument does not deviate from the function of the instrument.

After going through the small group trial stage and revising the test instruments based on the results of the small group trial analysis, the next stage is to conduct large group trials. This trial was carried out using 10 test items. Data on the validity of test instrument items in large group trials are presented in Table 6.

Table 6. Validity of Test Instrument Items in Large Group

Question Number	r_{count}	r_{table}	Validity Criteria
1	0.709		Valid
2	0.671		Valid
3	0.614		Valid
4	0.414		Valid
5	0.53	0.329	Valid
6	0.443		Valid
7	0.54		Valid
8	0.59		Valid
9	0.615		Valid
10	0.685		Valid

Based on the results of testing item validity using correlation *Product Moment* with value r_{table} at a significance level of 5%. The results of data analysis carried out on a large group with a total of 36 students (including 18 students from XI MIPA 4 and 18 students from XI MIPA 5) with a value of r_{table} of 0.329. Based on the basis for determining the validity of test items, if the correlation coefficient value r calculated > r table, then the test instrument items are significantly correlated (declared valid), so it can be concluded that the ten items that have been developed are declared valid (valid) in measuring students' physics problem solving skills.

Based on the results of reliability testing using Cronbach's Alpha, the results obtained were that in the small group trial the value was 0.848 with a very high reliability category, while in the large group trial the value reached 0.784 which was in the high category. If the reliability score is in the range of 0.70-0.90, then reliability is in the high category. This means that if the test instrument is tested several times, it will still give constant or the same results. The instrument is said to be reliable and acceptable if the Cronbach Alpha coefficient value is ≥ 0.7 with a confidence level of 95% (p: 0.05). Therefore, it can be concluded that the test instrument based on problem solving skills that was developed is feasible and accepted because it has a reliability value of r > 7.

Test instrument products that have been tested are then given a student response questionnaire to see students' responses to the test instruments that have been carried out. Analysis of student responses was carried out twice, including small groups and large groups. The results of the response analysis carried out in the small group obtained an average percentage of 81% and the results of testing the responses of students in the large group obtained an average response percentage of 81.32%. Based on the results of testing the responses of students in the two groups, a response value of > 80% was obtained, meaning that students showed a positive response to the test instrument based on physics problem solving skills and was suitable for use.

Discussion

The development of problem-solving skills-based test instruments in this study has gone through a series of phases and stages of the development of the 4-D model (Four D) by Thiagarajan starting from the define, design, develop and disseminate stages, resulting in a good test instrument product and has met the feasibility of a good instrument. This development research is not only limited to developing a test instrument product but also a test instrument product that can be used to measure students' physics problem solving skills. The test instrument products developed are in the form of descriptions on dynamic fluid material. The tests that have met the instrument's eligibility requirements can be used in the future, not only to measure problem solving skills, but also to measure higher order thinking skills like critical thinking (Elbyaly & Elfeky, 2023; Wider & Wider, 2023).

The feasibility of the test items developed has met the requirements of a good test instrument as a measuring tool if it has met the test requirements, namely having validity and reliability (Arikunto, 2016). This is emphasized by the view of Sudjana (2016) which states that an assessment is said to be good if the question has or meets two criteria, namely validity and reliability. The test instrument can be said to be good if it meets the requirements, one of which is content validity. The content validity of the test instrument needs to be done to assess the level of mastery of the content or content or certain material that should be mastered in accordance with the teaching objectives. The specific purpose of this validity is to obtain a good test instrument and the reason the instrument is said to be good or not good (Tanjung et al., 2021). Based on the calculation of the content validation test from the expert panel assessment, it shows that of the 10 items reviewed by 3 expert validators in the field of assessment, it was found that all ten questions were declared valid. This is evidenced by the average value of the percentage calculation reaching greater than 90% and after being matched with the validity criteria classified as a very valid category with minor revision requirements as suggested by the validator. The results of this validation show the instrument is a useful, appropriate and understandable for potential users (Kalkbrenner, 2021; Bernardo Gutiérrez et al., 2022).

The results of the percentage analysis of the three main aspects in the development of test instruments obtained an average for material aspects of 88% with a very good category, construction aspects of 85.5% with a very good category and language aspects of 85.25% with a very good category. Based on these percentages, overall, 10 items were obtained that had fulfilled the three aspects, namely material, construction and language with an average of the three aspects of 86.25%, which is in the very good category and can be continued in research testing. This is in accordance with the views of Tanjung and Nasution (2022) who state that test instruments that have met content validity can be used for further or future tests. The validation is based on a consistent research design across preferences, and applies state-of-the-art experimental techniques and transparent, quantitative criteria for test selection (Falk et al., 2023).

Based on the results of testing the validity of the items using the Product Moment correlation with the r table value at the 5% significance level. Item validity testing was carried out twice including tests in small groups and large groups. Based on the results of the validity test analysis in a small group with a total of 10 students and the r_{table} value shows 0.631, it is found that the ten items are declared to be in the valid category and can continue testing on a larger scale. The results of data analysis conducted in a large group with 36 students with a r_{table} value of 0.329. Based on the basis of determining the validity of the items if the correlation coefficient value r count> r table, then the test instrument items are significantly correlated (declared valid), so it can be concluded that the ten items that have been developed are declared valid (valid) in measuring students' physics problem solving skills. Additionally, the validity of the instruments was demonstrated, since the five factors postulated at a theoretical level for complex thinking (problem solving, critical analysis, metacognition, systematic analysis, and creativity) were assessed, and the results showed that the goodness of fit indexes met the criteria established in this field (Disabato et al., 2016; Tobón, & Luna-Nemecio, 2021).

Reliability is complementary to validity in understanding the accuracy of a measure. This means that an ultimately valid measure that exhibits low reliability is likely to produce at most low correlations with the target (Noble, Scheinost & Constable, 2019). Based on the results of reliability testing using Cronbach's Alpha, the results obtained in the small group trial obtained a value of 0.848 with a very high reliability category, while in the large group trial obtained a value of 0.784 which is in the high category. According to the test instrument eligibility criteria, if the reliability score is in the range of 0.70-0.90, the reliability is in the high category (Tanjung and Nasution, 2022). This means that if the test instrument is tested several times, it will still give the same or the same results. In Tanjung and Bakar's research (2019) stated that the reliability reached 0.763, meaning that the test instrument used had good reliability. This is also in line with the research of Tanjung and Rahma (2022) who obtained a reliability test result value of 0.80 with reliability in the very high category. The instrument is said to be reliable and acceptable if the Cronbach Alpha coefficient value is ≥ 0.7 with a 95% confidence level (p: 0.05) (Tanjung and Bakar, 2019; Hadisaputra et al., 2017). Therefore, it can be concluded that the test instrument based on problem solving skills developed can be feasible and acceptable because it has a reliability value of r> 7.

The test instrument that has been tested is then given a student response questionnaire to see the students' responses to the test instrument that has been done. Analysis of learner responses was carried out twice, including small groups and large groups. Based on the results of testing the response of students in both groups, a response value of> 80% was obtained, meaning that students showed a positive response to the physics problem solving skills-based test instrument developed. The results of the response analysis conducted in small groups obtained an average percentage of 81% with a very good category and the results of testing the response of students in large groups obtained an average percentage of positive student responses to the test instruments that have been developed of 81.32% which means that the instruments developed are very good and suitable for use. This is in line with research conducted by Tanjung and Dwiana (2019) where the results of student responses obtained were 75% in the good category, so that the problem-solving test instrument on vector material can be used as a material for measuring and assessing cognitive abilities for face-to-face and online learning.

This series of development research processes cannot be denied that there are several obstacles experienced by researchers. These obstacles include researchers having difficulty in finding reference sources and scientific articles that are the basis for developing test instruments. Some other obstacles faced by researchers when conducting research are that there are still many students who have difficulty and confusion in describing the stages of problem solving based on indicators of physics problem solving skills. This is because students are not accustomed to solving physics problems that require higher-level thinking skills, another obstacle is also the lack of time allocation. This also happens because the learning process is still carried out in a limited manner. Although there were several obstacles during the research process, the researchers have tried to minimize these obstacles. So it is hoped that this test instrument can fulfill the limitations of problem solving skills-based test instruments in the field and can be a reference for students in practicing their physics problem solving skills, especially for dynamic fluid material.

CONCLUSION

Based on the results and discussion of the preparation of test instruments based on problem solving skills on dynamic fluid materials which have been described previously, it can be concluded as follows: (1) Test instruments based on problem solving skills on dynamic fluid material have been successfully developed using Research and Development (R&D). (2) The test instrument in the form of an essay test consists of 10 questions. The test instrument developed has been declared feasible to use with the criteria that it has met expert validation and other supporting aspects such as item validity, construct validity, reliability, and student response tests.

REFERENCES

Ainun, F. P., Mawarni, H. S., Fauzah, N. N., & Raharja, R. M. (2023). Peran Pendidikan Sebagai Pondasi Utama Dalam Menyikapi Dekadensi Moral Pada Generasi Z. *Dewantara: Jurnal Pendidikan Sosial Humaniora*, 14–24.

Arikunto, S. (2016). *Prosedur Penelitian Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.

Arikunto, S. (2012). *Prosedur Penelitian Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.

Bernardo Gutiérrez, A. B., Esteban, M., Cervero Fernández-Castañón, A., Tuero Herrero, E., & Herrero Díez, F. J. (2022). Validation of the Early University Dropout Intentions Questionnaire (EUDIQ-R). *Journal of Higher Education Theory and Practice*.

Bozorova, H., & Shaxnoza, R. (2023). The Importance of Assessment, and Its Impact in The Learning Process. *Молодые ученые, 1*(6), 90–94.

Disabato, D. J., Goodman, F. R., Kashdan, T. B., Short, J. L., & Jarden, A. (2016). Different Types of Well-Being? A Cross-Cultural Examination of Hedonic and Eudaimonic Well-Being. *Psychological Assessment, 28*(5), 471.

Elbyaly, M. Y. H., & Elfeky, A. I. M. (2023). The Impact of Problem-Solving Programs in Developing Critical Thinking Skills. *European Chemical Bulletin, 12*, 6636–6642.

Falk, A., Becker, A., Dohmen, T., Huffman, D., & Sunde, U. (2023). The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences. *Management Science, 69*(4), 1935–1950.

Garcia, C., Argelagos, E., & Privado, J. (2021). Assessment of Higher Education Students' Information Problem-Solving Skills in Educational Sciences. *Information Development, 37*(3), 359–375.

Hadisaputra, S., Ihsan, M. S., Gunawan, & Ramdani, A. (2020). Pengembangan Perangkat Pembelajaran Kimia Didasarkan pada Model Blended Learning untuk Meningkatkan Kemampuan Berpikir Kritis Siswa. *Jurnal Fisika: Seri Konferensi, 1521*.

Heller, K., & Heller, P. (2010). *Pemecahan Masalah Kooperatif dalam Fisika Panduan Pengguna*. AS: Departemen Pendidikan.

Jayadi, A., Putri, D. H., & Johan, H. (2020). Identifikasi Pembekalan Keterampilan Abad 21 Pada Aspek Keterampilan Pemecahan Masalah Siswa SMA Kota Bengkulu dalam Mata Pelajaran Fisika. *Jurnal Kumparan Fisika*, 25–32.

Kara, F., & Celikler, D. (2015). Pengembangan Uji Prestasi: Validitas dan Reliabilitas.

Lewis, A., & Smith, D. (1993). Mendefinisikan Pemikiran Tingkat Tinggi. *Teori Menjadi Praktik, 32*(3), 131–137.

Kalkbrenner, M. T. (2021). A practical Guide to Instrument Development and Score Validation in The Social Sciences: The MEASURE Approach. *Practical Assessment, Research, and Evaluation, 26*(1), 1.

Kopnina, H. (2020). Education for the Future? Critical Evaluation of Education for Sustainable Development Goals. *The Journal of Environmental Education, 51*(4), 280–291.

Maison, D., Astalini, K., DA, S., & Perdana, R. (2020). Supporting Assessment in Education: E-Assessment Interest in Physics. *Universal Journal of Educational Research, 8*(1), 89–97.

Marlina, D., Sari, S. Y., & Rahim, F. R. (2024). Pengembangan Media Pembelajaran Interaktif Menggunakan Aplikasi Lectora untuk Meningkatkan Kemampuan Berpikir Kritis dan

Kreatif Siswa. *Jurnal Pengajaran Fisika Inovatif*, 23–36.

Mayasari, A., Asrizal, & Usmeldi. (2023). Effect Size Pengaruh Pembelajaran Berbasis SETS Terhadap Kemampuan Berpikir Kritis dan Hasil Belajar Siswa. *LENSA: Lentera Sains*, 67–76.

Mata-López, W. A., Juárez-Hernández, L. G., Tobón, S., & Montesinos-López, O. A. (2021). Analysis of Construct Validity and Reliability of Two Instruments to Evaluate Professional-Oriented Activities. *Revista Española de Orientación y Psicopedagogía (REOP), 32*(2).

Mohan, R. (2023). *Measurement, Evaluation and Assessment in Education*. PHI Learning Pvt. Ltd.

Noble, S., Scheinost, D., & Constable, R. T. (2019). A Decade of Test-Retest Reliability of Functional Connectivity: A Systematic Review and Meta-Analysis. *Neuroimage, 203*, 116157.

Purwanto. (2017). *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar.

Purwaningsih, E., Sari, S. P., Sari, A. M., & Suryadi, A. (2020). The Effect of STEM-PjBL and Discovery Learning on Improving Students' Problem-Solving Skills of Impulse and Momentum Topic. *Jurnal Pendidikan IPA Indonesia, 9*(4), 465–476.

Rahmawati, N., Nuryani, Y., Hamidah, I., & Rusdiana, D. (2020). Pengembangan Instrumen Tes Keterampilan Pemecahan Masalah pada Materi Kelistrikan dan Kemagnetan. *Conference Seminar Kontribusi Fisika*, 94–103.

Rahmi, M., Rahim, F. R., Sari, S. Y., & Murtiani, M. (2024). Development of a HOTS Indicator Availability Test Instrument for Instructional Media. *Journal of Innovative Physics Teaching, 2*(1), 78–88.

Rapih, S., & Sytardi. (2018). Perspektif Guru Sekolah Dasar Terhadap Higher Order Thinking Skills: Pemahaman, Penerapan, dan Hambatan. *Premier Education: Jurnal Pendidikan Dasar dan Pembelajaran*, 78–87.

Ratnasari, B., Haris, A., & Azis, A. (2021). Studi Kemampuan Berpikir Tingkat Tinggi Fisika di SMA. *Jurnal Sains dan Pendidikan Fisika (JSPF)*, 57–68.

Ridho, M. H., Wati, M., Misbah, M., & Mahtari, S. (2020). Validitas Bahan Ajar Gerak Melingkar Berbasis Authentic Learning di Lingkungan Lahan Basah untuk Melatih Keterampilan Pemecahan Masalah. *Journal of Teaching and Learning Physics*, 87–98.

Rusdianto, A. (2020). Pengembangan Instrumen Penilaian Berbasis Higher Order Thinking Skill (HOTS) Pelajaran Matematika Kelas IV Sekolah Dasar. *JP3D (Jurnal Pembelajaran dan Pengajaran Pendidikan Dasar)*, 2654–2870.

Sani, R. A. (2019). *Cara Membuat Soal HOTS*. Tangerang: Tira Smart.

Sani, R. A. (2019). *Strategi Belajar Mengajar*. Depok: Rajawali Pers.

Sari, I. P. M., Jatmiko, B., & Suprapto, N. (2023). Students' Physics Problem-Solving Skills in Daily Life Context: Between Confession and Fact. *Jurnal Penelitian Pendidikan IPA, 9*(1), 231–241.

Sudijono, A. (2008). *Pengantar Evaluasi Pendidikan*. Jakarta: Raja Grafindo Persada.

Sudjana, N. (2016). *Penilaian Hasil Proses Belajar Mengajar*. Bandung: PT Remaja Rosdakarya.

Sugiarto, & Farid, A. (2023). Literasi Digital Sebagai Jalan Penguatan Pendidikan Karakter di Era Society 5.0. *Cetta: Jurnal Ilmu Pendidikan*, 580–597.

Suprapto, W., Gustin, G., & Kariadi, D. (2023). Guru VS Media Sosial: Kontradiksi Peran Guru Di Era Global. *Jurnal Sustainable*, 148–158.

Tanjung, Y. I., & Dwiana, Y. A. (2019). Pengembangan Instrumen Tes Berbasis Critical Thinking Skill Pada Materi Gerak Lurus. *Jurnal Inovasi Pembelajaran Fisika*, 83–86.

Tanjung, Y. I., & Rahma, M. A. (2022). Pengembangan Instrumen Tes Berbasis Pengetahuan Faktual pada Materi Gelombang Bunyi di SMA. *Jurnal Ikatan Alumni Fisika Universitas Negeri Medan*, 1–7.

Tanjung, Y. I., & Bakar, A. (2019). Pengembangan Instrumen Tes Fisika Berdasarkan Dimensi Pengetahuan Konseptual Taksonomi Bloom Revisi. *Jurnal Pendidikan Fisika*, 141–148.

Tanjung, Y. I., & Nasution, I. R. (2022). The Development of Creative Thinking Test Instruments With Torrance Indicators on Direct Current Electricity Materials. *Jurnal Pendidikan Fisika Indonesia, 18*(2), 134–143.

Tanjung, Y. I., Wulandari, T., Festiyed, F., Yerimadesi, Y., & Ahda, Y. (2023). Development Analysis of Creative Thinking Test Instruments on Natural Science Materials. *Jurnal Pendidikan Fisika, 12*(1), 22.

Tanjung, Y. I., Wulandari, D., Bakar, A., & Ramadhani, I. (2021). The Development of Online Physics Test System at SMA CT Foundation Medan. *Journal of Physics: Conference Series, 1819*(1), 012054. IOP Publishing.

Tanjung, Y. I., Festiyed, F., Diliarosta, S. D., Jubaidah, J., & Sanjaya, A. P. S. (2024). Analysis of Students' Problem-Solving Skills Based on Heller indicator. *Pegem Journal of Education and Instruction, 14*(4), 421–433.

Tobón, S., & Luna-Nemecio, J. (2021). Complex Thinking and Sustainable Social Development: Validity and Reliability of The Complex-21 scale. *Sustainability, 13*(12), 6591.

Tugiah, T., & Jamilus, J. (2022). Pengembangan Pendidik Sebagai Sumber Daya Manusia untuk Mempersiapkan Generasi Milenial Menghadapi Era Digital. *Jurnal Sosial Teknologi*, 498–505.

Wandika, W., & Fitri, A. D. (2023). Analisis Besaran Pengaruh Model Pembelajaran Berbasis Masalah Terhadap Kemampuan HOTS dan Pemahaman Konseptual Siswa. *Jurnal Pengajaran Fisika Inovatif*, 40–49.

Widana, I. W. (2017). *Modul Penyusunan Higher Order Thinking Skill (HOTS)*. Jakarta: Direktorat Pembinaan SMA Direktorat Jenderal Pendidikan Dasar dan Menengah.

Wider, C., & Wider, W. (2023). Effects of Metacognitive Skills on Physics Problem-Solving Skills among Form Four Secondary School Students. *Journal of Baltic Science Education, 22*(2), 357–369.

Yanto, F., Festiyed, F., & Enjoni, E. (2021). Problem-Based Learning Model for Increasing Problem-Solving Skills in Physics Learning. *JIPF (Jurnal Ilmu Pendidikan Fisika, 6*(1), 53–65.