Jurnal Penelitian Pembelajaran Fisika
Vol. 16 Issue 2 – April 2025, p191-201
p-ISSN 2086-2407, e-ISSN 2549-886X
Available Online at http://journal2.upgris.ac.id/index.php/JP2F
DOI: 10.26877/jp2f.v16i2.1706

Development of Diagnostic Tests to Identify Students' Problem-Solving Difficulties in Static Fluid Material

Sri Wahyu Ningsih[1,2], Haratua Tiur Maria Silitonga[1] and Muhammad Musa Syarif Hidayatullah[1]
[1]Physics Education Study Program, Universitas Tanjungpura, Jl. Prof. Dr. H. Hadari Nawawi, Kota Pontianak

[2]E-mail: f1051211002@student.untan.ac.id

Received: 20 February 2025. Accepted: 29 April 2025. Published: 30 April 2025.

Abstract. The availability of diagnostic instruments specifically designed to identify students' difficulties in solving problems related to static fluids is still limited, even though this material often leads to misconceptions and conceptual difficulties. This article aims to present the results of the development of a diagnostic test instrument to identify students' problem-solving difficulties in the static fluid material. This study employs the R&D (Research and Development) method using the ADDIE model (Analysis, Design, Development, Implementation, Evaluation). The validation of the test instrument was assessed by two physics education lecturers and two physics teachers, covering content and language validation. The instrument validation results were analyzed using the Aiken index formula, obtaining a value of 0.83 with a high rating for content validation and a value of 0.93 with a high rating for language validation. The respondents in this study consisted of 95 students from grade XII. The test instrument provided consisted of essay questions, with each question containing two parts related to conceptual understanding and problem-solving ability. The analysis was conducted using Winsteps software. Based on the student response survey results regarding the diagnostic test, the average response percentage was 71%, which falls within the good criteria. The results of this study can be used as a reference for the development of instruments in physics, particularly static fluids, and help teachers better understand and address students' learning difficulties.

Keywords: diagnostic test, problem solving, static fluids

1. Introduction

Education is an effort made to prepare the younger generation to face the developments of the times in the global era. One of the important skills in the 21st century is the mastery of the 4Cs. The 4C skills include communication, collaboration, critical thinking and problem solving, and creative thinking [1]. One of the important components that students must possess in the development of the times is the ability in problem-solving [2]. Problem-solving skills are one of the higher-order thinking skills that students must possess. In the context of physics education, problem-solving skills need to be instilled in students so that they can become reliable problem solvers, both in physics and in everyday life [3].

Physics is a field of science that studies various phenomena occurring in nature. Until now, physics is still considered one of the subjects that is difficult for students. This is because most of the physics lessons are abstract and require a high level of reasoning, which affects students' understanding of concepts and problem-solving abilities. The results of the static fluid concept understanding test are still in the low category. The difficulty is caused by the numerous concepts in the material, making it hard for students to understand the entire concept [4].

Problem-solving is an effort to find solutions to mathematical problems by applying concepts or methods that have been mastered [5]. In physics learning, problem-solving is one of the important parts to apply physics concepts [6]. Problem-solving in physics education is often related to everyday life contexts. Problem-solving skills can help students solve everyday problems based on relevant theories and concepts [7]. Factors that influence the solving of physics problems are the knowledge structure possessed by the student solving the problem and the nature of the problem [8].

One way to determine students' problem-solving abilities is by administering a diagnostic test. Diagnostic tests need to be conducted to determine where the weaknesses and strengths of students lie in mastering the concepts of a part or the entire subject matter, as well as to identify emerging learning difficulties so that students' failures and successes can be known [9]. According to Inggit [10], diagnostic tests are a series of tests used to identify the weaknesses of students, so that the results can be used as a basis for providing follow-up actions in the form of appropriate handling that aligns with the weaknesses of the students. According to Rusilowati [11] diagnostic tests have the characteristics of diagnostic tests, namely: a) to detect learning difficulties, b) developed based on an analysis of the sources of difficulties, c) using supply response question forms (descriptive/short answer), d) if using selected response question forms, accompanied by reasons for the selection, and e) accompanied by follow-up plans, according to the identified difficulties.

One of the diagnostic tests that can be used is a diagnostic test in the form of an essay. According Sriyanti et al. [12], descriptive test is a test where the answers are given in the form of writing opinions based on the knowledge possessed. The knowledge measured by essay tests is high-level cognitive knowledge. Descriptive tests themselves have the advantage of being able to see the extent to which students understand their lessons from the answers they provide.

One of the problem-solving theories that can be used to determine students' problem-solving abilities is Heller's problem-solving theory. There are 5 stages of problem-solving according to Heller's Theory, including visualizing the problem, describing the problem, planning the solution, implementing the solution plan, and evaluating the solution [13]. If the problem-solving abilities of students have been identified, then the weaknesses in those abilities can be addressed by training them through appropriate learning.

Based on the results of the preliminary study, in the research conducted by Ringo et al. [14] namely analyzing students' problem-solving abilities in static fluid problems by providing descriptive problem-solving questions, which indicate that students' problem-solving abilities in static fluid material are still low, such as determining fluid depth in the sub-material of hydrostatic pressure. And the research conducted by Estianinur et al. [15] namely identifying students' problem-solving abilities in static fluid material, it was also found that 86.0% of students are classified as novice solvers in the sub-material of hydrostatic pressure, 88.4% in Pascal's Law, and 55.8% in Archimedes' Law. The novelty of this research lies in the use of Heller's theory to identify students' problem-solving difficulties based on the five stages of Heller's theory. Additionally, the developed test instruments are designed to include two types of questions: those related to the understanding of basic concepts in static fluids and those involving problem-solving using Heller's stages to observe students' problem-solving difficulties. The designed questions are conceptual in nature. The advantage of this research is to provide a comprehensive picture of the relationship between conceptual understanding and students' problem-solving abilities in static fluid material. Based on the initial observation conducted by distributing questionnaires to students and educators in the 11th grade at SMA Muhammadiyah 1 Pontianak, with a total of 38 students, it was found that many students still experience difficulties at the Heller stages, especially in describing the existing problems and explaining the issues.

Based on the problems that have been presented, the researcher intends to develop a diagnostic test in order to obtain a valid diagnostic test that can identify problem-solving difficulties in static fluid material. The main objective of this research is to develop a diagnostic test to identify students' problem-solving difficulties in static fluid material.

2. Method

This article is the result of research and development (R&D) using the ADDIE development model proposed by Robert M. Branch [16]. This model consists of five stages, namely: analyze, design, development, implementation, and evaluation.

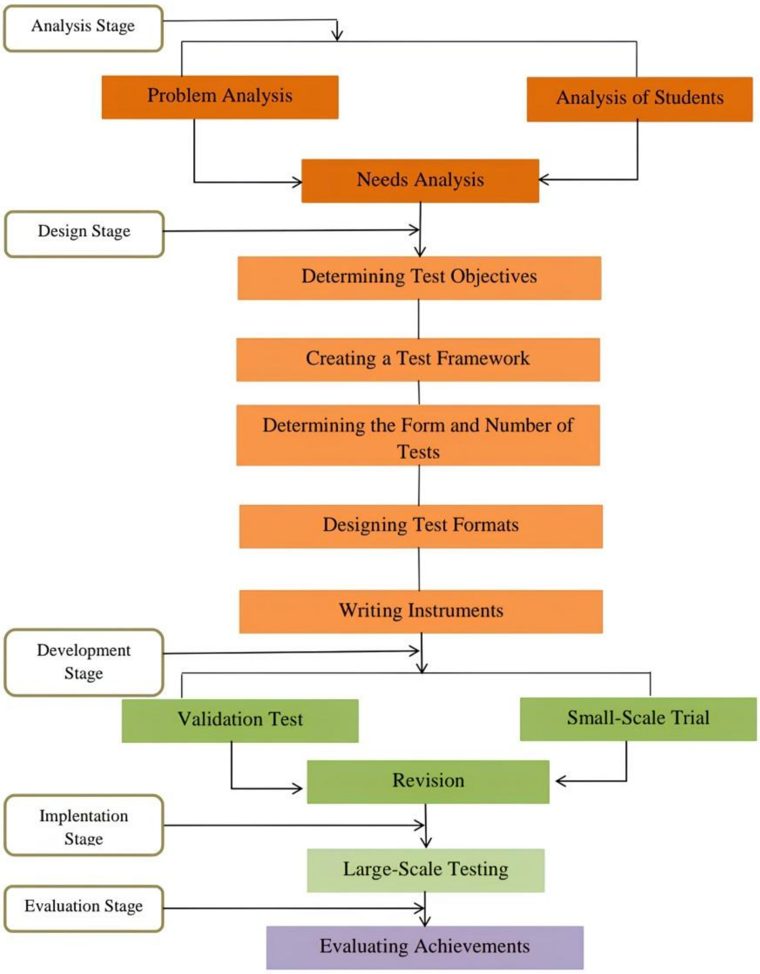

Figure 1. Development procedure.

The sample of this study consists of 95 twelfth-grade students who are enrolled in the Physics subject, focusing on static fluid material at SMA Negeri 5 Pontianak and SMA Santo Paulus Pontianak. Whereas, data collection was conducted during the Teaching and Learning Activities (KBM) of Physics in the odd semester. The test instrument used consists of problem-solving items in essay form for the topic of static fluids. The collected data were analyzed using the Rasch model with the help of Winsteps software to determine the validity, reliability, and difficulty level of the test items used.

Validation of the test instrument was also conducted with four validators, namely two physics education lecturers and two physics teachers, which included validation of content and language before the diagnostic test instrument was applied to the students. The validation results from the experts were analyzed using the Aiken index formula in equation 1

$$V = \frac{S}{[n(C-1)]}; S = \sum ni(r - l_0) \quad (1)$$

where V is the item validation index to measure the level of agreement among raters regarding the validity of an instrument. In this calculation, there are several categories or criteria (c) used to assess the instrument. Each category has certain limits called low categories (l_0), which indicate the level of invalidity of the instrument in that category. The evaluator (ni) is the number of evaluators who provide assessments in a certain category (r), while n refers to the total number of evaluators (raters). After the V index value is obtained, the value is classified based on its validity level to indicate the extent to which the instrument is considered suitable by the experts. The results of the average validation score calculation will be classified according to Table 1.

Table 1. Question instrument validation score criteria.

Value	Level of Criteria
0 – 0.4	Low
0.41- 0.8	Medium
0.81 – 1	High

To determine the level of student response to the product, the steps for processing the student response questionnaire data according to Riduwan [17] were used, which involve counting the number of respondents who chose four answer options based on the assessment scores for each item in Table 2.

Table 2. Student response score.

Score	Assessment criteria
4	Strongly Agree
3	Agree
2	Disagree
1	Strongly Disagree

The percentage of score acquisition for each item is calculated using the formula in equation 2.

$$P = \frac{\sum x}{\sum x_i} \times 100\% \quad (2)$$

This formula is used to measure the achievement level of students against the developed instrument. The percentage of score achievement (P) is obtained by dividing the total actual score obtained by the students on all items ($\sum x$) by the total maximum score that can be achieved ($\sum x_i$), and then multiplying by 100%. The conversion into percentage form aims to facilitate the process of interpreting and classifying the level of student achievement based on the established criteria. The results obtained are interpreted through Table 3.

Table 3. Response interpretation.

Percentage of Assessment	Interpretation
0% - 20%	Very Bad
21% - 40%	Not Good Enough
41% - 60%	Good Enough
61% - 80%	Good
81% - 100%	Very Good

3. Results and Discussion

The development that has been carried out is the result of applying the development steps of the ADDIE model with the stages of analysis, design, development, implementation, and evaluation. The development product is in the form of a diagnostic test instrument to identify students' difficulties in problem-solving on static fluid material.

3.1. *Analysis Stage*

An initial observation was conducted by distributing a needs analysis questionnaire. Needs analysis is conducted to develop tests that align with the needs of the learners. This analysis stage is conducted by collecting student needs questionnaires. This questionnaire is presented in the form of a Google Form with a total of 11 statements for the student needs questionnaire. The collected student needs survey data amounted to 38 people. Where students have difficulty in describing problems in physics concepts, namely students write known and unknown variables in physics symbols and use the International

System of Units (SI). The next difficulty is at the stage of describing the problem, where students struggle to describe the problem in the question and to write down the known variables. The next difficulty is in the evaluation stage, where students struggle to conclude that the answer obtained is correct and reasonable. The next difficulty is in the planning solution stage, where students struggle to determine the mathematical formulation to be used in solving the problem. This can occur because students tend to directly use formulas they have memorized from frequently given example problems. And finally, the difficulty in executing the solution plan, this difficulty is caused by the students' lack of calculation skills.

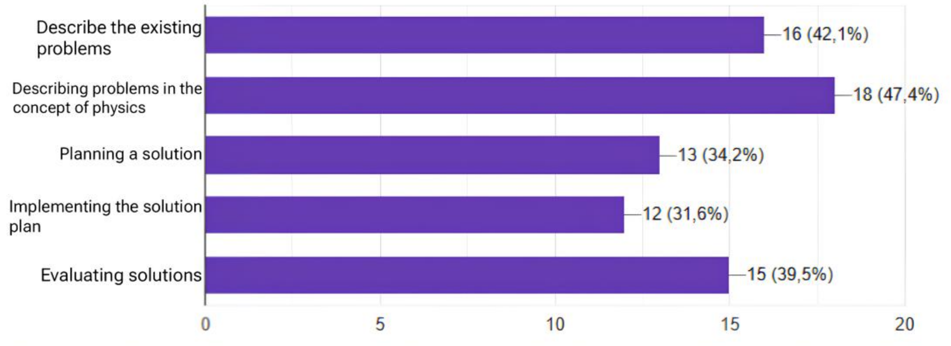

Figure 2. Diagram of student difficulties in problem solving.

Based on Figure 2, the results of the needs analysis questionnaire from 38 students indicate that they experience difficulties in solving problems at each stage of the Heller problem-solving process. Based on the results of this needs analysis, it shows that students are accustomed to solving physics problems by directly using mathematical equations without conducting an analysis, guessing the formulas used, and memorizing example problems that have been solved to tackle other problems [18]. However, educators also stated that most students in the class experienced difficulties in the stage of planning solutions due to a lack of calculation skills. And based on the diagram in Figure 1, 47.4% of the students chose to have difficulty in the second stage, which is describing the problem in the concept of physics.

3.2. *Design Stage*

At the design stage, instrument development refers to Heller's problem-solving theory, based on Heller's solution steps which consist of five stages including Visualize the problem, Physics description, Plan a solution, Execute the plan, and Check and evaluate [13]. In addition to problem-solving questions, there are also conceptual understanding questions aimed at measuring students' understanding of fundamental concepts in physics. Diagnostic tests of physics concept understanding can also help teachers adjust their teaching methods to meet the individual learning needs of students [19]. And the use of Bloom's taxonomy as a basis for determining cognitive levels and formulating question indicators. Bloom's taxonomy provides guidance for formulating clear and measurable learning objectives. This framework helps educators choose teaching methods that are appropriate for the students' skill levels. Additionally, Bloom's Taxonomy also facilitates the development of evaluation tools that are relevant to the expected learning outcomes [20].

This stage involves the initial design of the instrument by creating a diagnostic test instrument, which includes determining the title/subtitle of the diagnostic test instrument, then designing it in accordance with the learning outcomes of static fluid material and the flow of learning objectives, preparing the test blueprint, determining the form of the test and the number of questions, as well as designing the validator and response questionnaire according to the characteristics of the students. Thus, a blueprint for the diagnostic test instrument was produced to identify problem-solving difficulties in static fluid material, consisting of 5 questions.

Table 4. Test instrument framework.

No	Material	Indicator Question	Question Number
1	Hydrostatic Law	Students can determine the total pressure experienced by a diver by accurately summing the hydrostatic pressure and the external air pressure.	1
		Students can prove that the deeper the water in the tank, the greater the hydrostatic pressure generated at the bottom of the tank, and understand how the principles of hydrostatic law work correctly.	5
2	Archimedes' Principle	Students can relate the principle of Archimedes' Law to the magnitude of force, weight of the object, and mass to solve problems accurately.	2
		Students can accurately relate the concept of Archimedes' principle and how the balance of forces works on a person partially floating in a liquid.	3
3	Pascal's Law	Students can correctly compare the force required to lift a sedan and a minivan.	4

The questions designed consist of 5 questions, specifically on the subtopics of Hydrostatic Law, Archimedes' Law, and Pascal's Law. Each question contains two parts, the first related to understanding the basic concepts of static fluid material and the second related to solving using the 5 stages. Heller to observe the difficulties in problem-solving among students, namely the stages of visualize the problem, physics description, plan a solution, execute the plan, and check and evaluate.

3.3. *Development Stage*

After the diagnostic test instrument with 8 questions was designed, the diagnostic test was then validated by subject matter and language experts, namely 2 physics education lecturers and 2 physics teachers. The aspects evaluated by content experts include content appropriateness, material appropriateness, and construction appropriateness. Validation aims to ensure that the developed diagnostic test aligns with the theory of static fluids. Then, teacher validation aims to ensure that the developed diagnostic test aligns with the static fluid material taught in schools. Language validation has assessment criteria, namely using language that adheres to standard language rules, is communicative, does not lead to multiple interpretations, and has clear instructions. This validation aims to ensure that the developed diagnostic test adheres to the proper and correct rules of Indonesian language writing [21].

After the validation process was carried out, several inputs were obtained that became the basis for revising the design of the diagnostic test instrument used to identify students' problem-solving difficulties. The revision was carried out based on the assessment results by subject matter experts and language experts. The revisions cover the material aspect, particularly adjusting the use of numbers in the questions to better align with real-life contexts. Additionally, the term "speed" in one of the questions was changed to "speedboat" to enhance clarity and relevance of meaning. This revision also considers the importance of maintaining the conceptual character of the question.

The validation results involved submitting a questionnaire to 4 validators, namely 2 physics education lecturers and 2 physics teachers. Comments and suggestions from the validators were used as the basis for improvements before the diagnostic test instrument was trialed with the students. The results of the validation questionnaire above are the validation results from subject matter experts and language experts. The data obtained from the validation of the assessment instrument by the validators were then calculated using the Aiken index formula [22].

Table 5. Aiken index validation results.

No	Specialist	Index Validation Result	Category
1.	Material	0,83	High
2.	Language	0,93	High

Based on Table 5, the results of the data analysis from the validators were obtained using the Aiken index formula. The score for the subject matter expert was 0.83 with a high rating, and the language expert was 0.93 with a high rating. Research conducted by Novithania [23] also showed similar validation results, with the Aiken index (V) value falling into the valid category with a high description. Thus, this product can be used by students with improvements. To obtain an instrument that is truly valid, improvements or revisions need to be made based on the suggestions from the validator [24].

After obtaining valid results from the trials conducted by subject matter and language experts, trials were conducted on the students. The trial was conducted with 3 eleventh-grade students who are enrolled in the physics subject. To obtain data related to the difficulties experienced by students in completing the diagnostic test, a response questionnaire was distributed. The scores given by the 3 respondents were calculated using a Likert scale.

Table 6. Results of the small group likert scale.

No	Evaluator	Percentage(%)	Category
1.	First	80,36	Very Good
2.	Second	69,64	Good
3.	Third	62,50	Good

Based on table 6, the results of the data analysis from 3 respondents evaluated using the Likert scale were obtained. Respondent 1 scored 80.36% with a very good category, respondent 2 scored 69.64% with a good category, and respondent 3 scored 62.50% with a good category. Thus, this product can be used by educators for students with improvements based on the feedback and suggestions provided.

3.4. *Implementation Stage*

The implementation stage is the field trial phase of the test instrument that has passed the development stage, which has been revised based on notes from subject matter and language experts. The questions that have been deemed suitable are then given to the students for testing. The product trial is conducted by testing the revised product on students who have studied the material on static fluids. The implementation stage aims to determine the response or feedback from students regarding the diagnostic test developed by the researcher. At this stage, a large-scale trial was conducted with 95 students from the 12th grade of Senior High School (SMA) who are taking physics. Winsteps with the Rasch model approach to determine the reliability of test items, the difficulty level of test items, and the discrimination power.

The reliability test of the research instrument is to determine whether the test tool used to collect research data is reliable. The term "reliability" is used to describe how consistent the measurement results are when conducted two or more times with the same instrument and measuring tool [25]. In the Rasch model, according to Sumintono [26] the criteria for determining the values of Item Reliability and Person Reliability are based on the criteria of exceptional, very good, good, sufficient, and weak. Based on the results of the reliability analysis of the diagnostic test items using the Rasch model, it is detailed in Table 7.

Table 7. Reliability analysis results.

Reliability	Result	Category
Person reliability	0,69	Enough
Item reliability	0,99	Special
Reliability alpha cronbach's	0,72	Good

Based on Table 7, the Cronbach's alpha value of 0.72 indicates good interaction. Next, the person reliability value of 0.69 indicates a sufficient consistency in students' answers, categorized as sufficient, while the Item Reliability value of 0.99 falls into the excellent category, meaning the quality of the test items in the instrument has a good aspect of reliability. Reliability is the interaction between the individual and the test items as a whole, which can be measured using the Cronbach's alpha value. The individual reliability score indicates the consistency of students' answers, and the item reliability score indicates the quality of the questions [27].

In Rasch modeling, the difficulty level of test items is categorized based on the logit Measure and divided into five categories: very easy, easy, moderate, difficult, and very difficult [22]. In this study, there are 95 respondents with a total of 30 items. The results of the item analysis show that the difficulty level criteria are divided into five categories. Items are categorized as easy if they have a b value close to -2.00 logit, categorized as moderate if -1.00 logit < b < +1.00 logit, and categorized as difficult if the b value is close to +2.00 logit. Next, items with a value of b > +2.00 logit fall into the very difficult category, and items with a value of b < -2.00 logit fall into the very easy category [28]. Which can be seen in the Rasch modeling results in Figure 3.

Person: REAL SEP.: 1.50 REL.: .69 ... Item: REAL SEP.: 6.76 REL.: .98

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
14	36	95	4.45	.22	1.01	.13	.98	-.09	.47	.36	74.7	68.2	S14
8	46	95	3.98	.21	1.25	2.22	1.20	1.72	.63	.37	70.5	64.4	S8
20	72	95	2.86	.21	1.55	3.54	1.57	3.47	.59	.37	48.4	68.0	S20
2	84	95	2.33	.21	2.21	5.86	2.27	5.87	.75	.36	30.5	72.3	S2
18	88	95	2.15	.21	.22	-7.20	.20	-7.35	.50	.36	92.6	73.1	S18
12	89	95	2.10	.21	.23	-7.10	.21	-7.24	.44	.36	93.7	73.3	S12
6	91	95	2.01	.21	.19	-7.63	.18	-7.74	.39	.36	95.8	73.6	S6
24	95	95	1.82	.22	.15	-8.39	.14	-8.46	.00	.36	100.0	73.9	S24
30	95	95	1.82	.22	.15	-8.39	.14	-8.46	.00	.36	100.0	73.9	S30
26	100	95	1.59	.22	2.83	7.54	2.89	7.51	.66	.36	17.9	73.7	S26
10	102	95	1.50	.21	.26	-6.51	.25	-6.51	.42	.36	96.8	73.4	S10
11	109	95	1.18	.21	.73	-1.83	.65	-2.37	.58	.37	91.6	72.0	S11
16	109	95	1.18	.21	.43	-4.65	.41	-4.70	.45	.37	89.5	72.0	S16
17	113	95	1.00	.21	1.09	.61	1.00	.06	.53	.37	88.4	70.8	S17

Figure 3. Item statistics from winsteps.

Based on Figure 3, the analysis of item difficulty is seen from the measure count column, resulting in the analysis of the difficulty level of the diagnostic test items modeled by Rasch, as detailed in Table 8.

Table 8. Results of the difficulty level analysis of the questions.

No	Measure	Category
1.	5 (-3.58), 15 (-2.19), 19 (-2.69), 21 (-2.91), 22 (-2.91), 23 (-11.70), 27 (-2.91), 28 (-2.91), 29 (-11.70)	Very Easy
2.	1 (-1.75), 3 (-2.67), 4 (-0.51), 7 (-0.96), 9 (-0.73), 13 (-1.21), 25 (-1.93)	Easy
3.	10 (1.50), 11 (1.18), 16 (1.18), 17 (1.00), 24 (1.82), 26 (1.59), 30 (1.82)	Difficult
4.	2 (2.33), 6 (2.01), 8 (3.98), 12 (2.10), 14 (4.45), 18 (2.15), 20(2.86)	Very Difficult

In the analysis of item difficulty using the Rasch model, 9 items were categorized as very easy, 7 items as easy, 7 items as difficult, and 7 items as very difficult. In the concept understanding items, questions number 1, 7, 13, and 25 fall into the easy category, while question number 19 falls into the very easy category. In the Heller stage 1 question items, which involve describing the problem, question items number 2, 8, 14, and 20 fall into the very difficult category, while question item 26 falls into the difficult category. In the Heller stage 2 question items, which involve describing the problem in terms of physics concepts, question items number 3 and 9 fall into the easy category, while question items number 15, 21, and 27 fall into the very easy category. In the Heller stage 3 question items, which involve designing solutions, question item number 4 falls into the easy category, numbers 10 and 16 fall into the difficult category, and then numbers 22 and 28 fall into the very easy category.In the Heller stage 4 question items, which involve implementing the solution plan, items number 5, 23, and 29 fall into the very easy category, while items number 11 and 17 fall into the difficult category. In the Heller stage 5 question, which is evaluating the solution, questions number 6, 12, 18, and 20 fall into the very difficult category, while question number 30 falls into the difficult category.

The discriminating power of a test item is the ability of the item to differentiate between students who can answer the item or students with a high level of ability and students who have a low ability to answer the item. The results of the calculation of the item discrimination index can generally be categorized into five categories as shown in Table 9.

Table 9. Interpretation of discriminative power.

Differentiating Factor	Interpretation	Question Item
Negative	Drop	1 item (15)
0.00 – 0.20	Bad	19 item (3,4,6,7,9,10,11,13,16,29,21,22,23,24,25,27,28,29,30)
0.21 – 0.40	Enough	6 butir (1,5,12,14,17,18)
0.41 – 0.70	Good	2 item (8,20)
0.71 – 1.00	Very Good	2 item (2,26)

Based on Table 9, it can be seen that the results of the analysis of the question's discrimination power through calculations using equation 3

$$DP = \frac{B_A}{J_A} = \frac{B_B}{J_B} = P_A - P_B \quad (3)$$

where indicates that most of the question discrimination power falls into the poor category. The question's discrimination power that falls into the drop category is 1 item at number 15 with a value of -0.08, which is at the second Heller stage describing problems in physics concepts; the poor category has 19 items with values of 0.00 – 0.18; the sufficient category has 6 items with values of 0.21 – 0.38; the good category has two items with values of 0.50 – 0.58; and the very good category has 2 items with values of 0.79-0.85.

To obtain data related to the difficulties experienced by students in completing the diagnostic test, a response questionnaire was distributed. The instrument used is a closed questionnaire for student responses using a Likert scale with four assessment criteria, namely Strongly Agree; Agree; Disagree; and Strongly Disagree. The questionnaire consists of 14 statements, with 7 positive statements and 7 negative statements. There are 3 indicators in the questionnaire, namely: (1) The ability of students to understand the diagnostic test; (2) The attitude and perception of students towards the diagnostic test; (3) The ability of students to complete the test [29]. Based on the results of the student response questionnaire to the diagnostic test for identifying students' problem-solving difficulties in static fluid material, the percentage results obtained were 73% for students' ability to understand the diagnostic test, 70% for students' attitudes and perceptions towards the diagnostic test, and 71% for students' ability to complete the test. The average percentage score obtained from the student response questionnaire was 71%, categorized as good.

3.5. *Evaluation Stage*

At this stage, an evaluation is conducted for each phase in the development of the ADDIE model. First, at the analysis stage, the evaluation is conducted based on the results of the needs analysis questionnaire filled out by the students. The results show that students' ability to solve problems is still relatively low, because students are accustomed to solving physics problems by directly using mathematical equations without conducting an analysis. Second, at the design stage, the evaluation was conducted based on comments and suggestions from experts. Fourth, at the implementation stage, evaluation was conducted through the application of diagnostic test instruments to 95 respondents, which resulted in recommendations to improve several sentences in the questions that were difficult to understand. The evaluation conducted at each stage of the ADDIE model aims to ensure that the developed diagnostic test product on static fluid material is appropriate and suitable for identifying students' difficulties in problem-solving.

4. Conclusion

The results of the research and development indicate that a test instrument has been obtained to identify students' problem-solving difficulties in static fluid material. Based on the results of data analysis from the validators calculated using the Aiken index formula. The score for the subject matter expert was 0.83 with a high rating, and for the language expert, it was 0.93 with a high rating. Next, the analysis of item reliability obtained a Cronbach's alpha reliability value of 0.72, categorized as good, and an Item Reliability value of 0.99, categorized as excellent. The results of the difficulty level analysis based on the Rasch model show that 9 items fall into the very easy category, 7 items into the easy category, 7 items into the difficult category, and 7 items into the very difficult category. And from the analysis of the measurement instrument, the question discrimination power falls into the drop category with 1 item, the poor category with 19 items, the sufficient category with 6 items, the good category with 2 items, and the very good category with 2 items.

Acknowledgment

The author expresses gratitude to the FKIP Universitas Tanjungpura (UNTAN) for supporting the implementation of MBKM Research in 2024. Thank you to the Physics Education Study Program at FKIP UNTAN for all the support and facilities provided.

References

[1] Bura E L, Yeyen Y Y, Suban M E and Wajong A 2024 Analisis Kemampuan Pemecahan Masalah Peserta Didik pada Materi Fluida Statis Berbantuan 3D Pageflip Professional *Journal on Education* **06** 15871–8

[2] Gunur B, Parinters Makur A and Hendrice Ramda A 2018 Hubungan Antara Kemampuan Numerik Dengan Kemampuan Pemecahan Masalah Matematis Siswa Di Pedesaan *MaPan* **6** 148–60

[3] Taqwa M R A, Purwaningsih E and Sulur S 2020 Pengembangan Instrumen Tes Kemampuan Pemecahan Masalah Mahasiswa pada Topik Usaha dan Energi *Jurnal Penelitian Pembelajaran Fisika* **11** 149–56

[4] Amalissholeh N, Sutrio S, Rokhmat J and Gunada I W 2023 Analisis Kesulitan Belajar Peserta Didik pada Pembelajaran Fisika di SMAN 1 Kediri *Empiricism Journal* **4** 356–64

[5] Fauziah N, Roza Y and Maimunah M 2022 Kemampuan Matematis Pemecahan Masalah Siswa dalam Penyelesaian Soal Tipe Numerasi AKM *Jurnal Cendekia : Jurnal Pendidikan Matematika* **6** 3241–50

[6] Asuri A R, Suherman A and Darman D R 2021 Penerapan Model Problem Based Learning (PBL) Berbantu Mind Mapping dalam Pembelajaran Fisika untuk Meningkatkan Kemampuan Pemecahan Masalah pada Materi Usaha dan Energi *Jurnal Penelitian Pembelajaran Fisika* **12** 22–8

[7] Haryati E D, Sitompul S S and Hamdani 2019 Analisis Kemampuan Pemecahan Masalah Peserta Didik Smk Hidayatul Mubtadi ’ien Pada Materi Gerak Lurus *Jurnal Pendidikan dan Pembelajaran* **10** 1–11

[8] Sujarwanto E, Hidayat A and Wartono 2014 Kemampuan pemecahan masalah fisika pada modeling instruction pada siswa sma kelas xi *Jurnal Pendidikan IPA Indonesia* **3** 65–78

[9] Wulandari F 2019 Miskonsepsi Siswa Tentang Suhu Dan Kalor Menggunakan Tes Diagnostik Di Sma Negeri 1 Sejangkung Artikel *J Chem Inf Model* **53** 1689–99

[10] Inggit S M, Liliawati W and Suryana I 2021 Identifikasi Miskonsepsi dan Penyebabnya Menggunakan Instrumen Five-Tier Fluid Static Test (5TFST) pada Peserta Didik Kelas XI Sekolah Menengah Atas *Journal of Teaching and Learning Physics* **6** 49–68

[11] Rusilowati A 2015 Pengembangan Tes Diagnostik Sebagai Alat Evaluasi Kesulitan Belajar Fisika *Prosiding Seminar Nasional Fisika dan Pendidikan Fisika* **6** 1–10

[12] Sriyanti A, Mania S and A N H 2019 Pengembangan Instrumen Tes Diagnostik Berbentuk Uraian Untuk Mengidentifikasi Pemahaman Konsep Matematika Wajib Siswa Man 1 Makassar *De Fermat : Jurnal Pendidikan Matematika* **2** 57–69

[13] Lestari K, Maria S H T and Mahmuda D 2019 Penerapan Penyelesaian Masalah Heller Untuk Meningkatkan Kemampuan Menyelesaikan Soal Materi Gera Lurus *Jurnal Pendidikan dan Pembelajaran Khatulistiwa* **8** 2–9
[14] Ringo E S, Kusairi S and Latifah E 2019 Profil Kemampuan Pemecahan Masalah Siswa SMA pada Materi Fluida Statis *Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan* **4** 178
[15] Estianinur E, Parno P and Latifah E 2020 Identifikasi Kemampuan Pemecahan Masalah Siswa Materi Fluida Statis *Briliant: Jurnal Riset dan Konseptual* **5** 477
[16] Branch R M and Varank İ 2009 *Instructional design: The ADDIE approach* vol 722 (Springer)
[17] Riduwan 2016 *Dasar-Dasar Statistika* (14th ed) (Bandung: Alfabeta CV)
[18] Azizah, R; Yuliati, L; Latifah E 2015 Kesulitan Pemecahan Masalah Fisika Pada Siswa SMA *Jurnal Penelitian Fisika Dan Aplikasinya* **5** 44–50
[19] Wahyono U, Hermanto I M, Nurhayati N, Samatowa L, Mohamad W M and Maharani N L S 2023 Pengembangan Tes Diagnostik Untuk Mengidentifikasi Pemahaman Konsep Siswa Sma Pada Pokok Bahasan Gelombang Bunyi *Jambura Physics Journal* **5** 67–79
[20] Marta M A, Purnomo D, Islam U, Imam N and Padang B 2025 Konsep Taksonomi Bloom dalam Desain Pembelajaran **3**
[21] Prabaningtias, Dea, Arsi; Silitonga, Haratua, Tiur , Maria; Mahmudah D 2018 Pengembangan Tes Diagnostik Four Tier Menggunakan Aplikasi Google Form Pada Materi Fluida Statis SMA Pontianak *Jurnal Pendidikan dan Pembelajaran Khatulistiwa* **7** 1–10
[22] Z R P, Sari R, Jumadi J and Ariswan A 2020 Pengembangan dan Validasi Instrumen Tes untuk Mengukur Keterampilan Menyelesaikan Masalah Peserta Didik SMA pada Pelajaran Fisika *Jurnal Penelitian Pembelajaran Fisika* **11** 17–26
[23] Novithania C A, Maria H T, Tanjungpura U, Artikel I, Tes P, Siswa K B and Education J 2025 Pengembangan tes diagnostik testlet untuk mendeteksi kesulitan belajar pada materi listrik dinamis **13** 121–6
[24] Retnawati H 2016 *Analisis kuantitatif instrumen penelitian (Panduan Peneliti, Mahasiswa, dan Psikometrian)* (Yogyakarta: Parama publishing)
[25] Ono S 2020 Uji Validitas dan Reliabilitas Alat Ukur SG Posture Evaluation *Jurnal Keterapian Fisik* **5** 55–61
[26] B, Sumintono; W W 2015 *Aplikasi Pemodelan Rasch pada Assessment Pendidikan* (Cimahi : Trim Komunikata)
[27] Triyanti, Asih; Yuniarti A A 2025 Analisis Instrumen tes Pengambilan Keputusan Taksonomi Presseisen dan Abilitas Peserta Didik Menggunakan Model Rasch **11** 46–58
[28] Kumalasari E D and Mahmudi I 2024 *Analisis Pemodelan Rasch Pada Asesmen Pendidikan* (Banyumas: PT. Pena Persada Kerta Utama)
[29] Mardianto Y, Azis L A and Amelia R 2022 Menganalisis Respon Siswa Terhadap Pembelajaran Materi Perbandingan Dan Skala Menggunakan Pendekatan Kontekstual *JPMI (Jurnal Pembelajaran Matematika Inovatif)* **5** 1313–22