

[Research Article]

DEVELOPMENT OF STATIC FLUIDS ASSESSMENT USING EVIDENCE-CENTERED DESIGN (ECD) APPROACH

Libertus Elusandy¹, Judyanto Sirait¹, M. Musa Syarif. H¹

¹Department of Physics Education, Tanjungpura University, Pontianak, 78124 Indonesia
E-mail: F1051211017@student.untan.ac.id

DOI: <http://dx.doi.org/10.15575/jotalp.v11i1.54507>

Received: 28 January 2026 Accepted: 10 February 2025 Published: 28 February 2026

ABSTRACT

This research aims to create a static fluid assessment that physics teachers can use, utilizing the Evidence-Centered Design approach. Evidence-Centered Design (ECD) is a framework for designing educational assessments that uses a systematic approach to design, develop, and evaluate assessments focused on evidence collection. ECD consists of three main components: claims, evidence, and questions. This research focuses on static fluids as its topic. Research and development (R&D) adapts the ADDIE model (analysis, design, development, implementation, and evaluation) to create the assessment. Analysis is conducted through several stages, namely problem analysis, curriculum analysis, student analysis, and material analysis. The design stage is carried out by determining the assessment objectives, determining the form and number of assessments, designing the assessment components, and writing the assessments. The development stage involves three steps: expert validation, one-to-one trials, and small-scale trials. The results of the research and development of the ECD assessment are deemed feasible based on the results: 1) Validators stated that the instrument falls into the very feasible category with an average Aiken index of 0.87, categorized as very high; 2) one-to-one trial respondents stated that the instrument was very good and the time given was sufficient; and 3) a small-scale trial with a sample of 10 students found that the instrument was very appropriate and the time given was sufficient. Through assessment based on the Evidence-Centered Design (ECD) approach, it is hoped that it can provide guidance to teachers in designing assessments that are appropriate and effective in evaluating each student's understanding

Keywords: Assesment Development, Evidence-Centered Design, Static Fluids

How to cite: Elusandy, L., Sirait, J., and Syarif, M. M. H. (2026). Development of Static Fluids Assessment Using Evidence-Centered Design (ECD) Approach, *Journal of Teaching and Learning Physics* 11 (1), 30-41. DOI: <http://dx.doi.org/10.15575/jotalp.v11i1.54507>



1. INTRODUCTION

Physics is a branch of natural sciences (IPA) that studies natural phenomena in a structured and systematic manner (Nur'aini et al., 2020). According to Salassa et al. (2023), the goal of physics education is to provide students with an understanding so that they can master physics concepts and their relationships, enabling them to solve problems related to everyday life. Mastery of physics should not only make students merely know or memorize concepts but also understand and be able to apply that knowledge in real life (Yuwarti et al., 2017). To achieve this goal, a learning assessment is needed that can accurately measure students' abilities in accordance with the learning objectives designed by the teacher.

Assessment is essentially the process of collecting information and making decisions based on data obtained through measuring student learning outcomes, whether through tests or non-test instruments (Maulani et al., 2020). Kellaghan & Greaney (2001) also stated that assessment is a process of collecting, interpreting, and utilizing evidence related to an individual's knowledge, attitudes, and skills. In this context, assessment becomes an inseparable part of the learning process and helps educators in planning learning programs to obtain information objectively, continuously, and comprehensively regarding the process and outcomes of student learning. The results of the assessment are then used to determine the next steps or actions (Kurniawati & Sukardiyono, 2018). In other words, assessment is not only aimed at achieving short-term goals or limited to one aspect but must be comprehensive and cover various aspects, such as cognitive, affective, and psychomotor (Sukmawa et al., 2019)

The assessments designed by teachers often do not fully align with the established learning objectives. Research by Indriani et al. (2023) revealed that some teachers still tend to create assessments oriented towards Low Order Thinking Skills (LOTS), such as multiple-choice

questions, essays, and short answer questions. In those assessments, teachers place more emphasis on the competencies of remembering, understanding, and applying. However, current needs are more focused on mastering higher-order competencies like analyzing, evaluating, and applying, typically found in Higher Order Thinking Skills (HOTS) assessments. Unfortunately, the assessments created are less capable of encouraging students to think systematically, critically, logically, and analytically. This is due to the tendency of teachers to focus on rote memorization of facts, objective questions (true/false), and measuring mastery of material alone, without paying attention to the development of higher-order thinking skills, creativity in problem-solving, and comprehensive student competencies.

In science education, the main goal of teaching and learning is to help students develop a deep understanding of essential scientific principles and concepts (Bao & Koenig, 2019). One of the competencies that students need to possess in science education is higher-order thinking skills, such as analysis, evaluation, synthesis, and creativity in problem-solving. This ability allows students not only to understand scientific concepts theoretically but also to apply them in real situations to solve complex problems, design innovative solutions, and make decisions based on scientific evidence (Kistiono, 2019). Therefore, a solution is needed that can assist teachers in designing appropriate assessment questions to measure students' understanding during the learning process and develop higher-order thinking skills and problem-solving abilities required today, using the Evidence-Centered Design (ECD) approach.

Evidence-Centered Design (ECD) is a framework for designing educational assessments that uses a systematic approach to design, develop, and evaluate assessments focused on evidence collection (Mislevy & Riconscente, 2005). According to Mislevy & Haertel (2006), ECD is designed to ensure that assessments are developed systematically to collect relevant

evidence related to students' abilities, thereby providing a clear picture of the extent of their understanding. Based on the findings of (Newton et al., 2021), it is revealed that ECD not only assesses students' understanding in depth but also provides important information to teachers regarding students' progress in achieving learning objectives. According to Bechard et al. (2019), ECD helps teachers design evaluation instruments that require students to think critically and make evidence-based decisions. Thus, this approach provides students with the opportunity to develop essential skills for solving complex problems. This process involves the use of higher-order thinking skills that play a crucial role in 21st-century learning.

According to Arieli-Attali et al. (2019), in the ECD framework, there are several components, namely claims, evidence, and questions. A claim articulates the intended measurement or evaluation of the student. Evidence is data or information collected from students' answers to support or assess the truth of a claim, while assessment questions or tasks are instruments or tools designed to gather evidence from students (Zieky, 2014). Therefore, these three components must be integrated into the design of ECD-based assessments to ensure effective assessments in evaluating each student's learning achievements. (Lane, 2015).

In formulating learning objectives that are part of the claim component in ECD. The development of assessments is integrated with the element of scientific practice (SP). Scientific Practice (SP) is a competency that focuses on enhancing students' skills and knowledge to foster a profound scientific understanding (Duschl & Bybee, 2014). Integrating SP into assessments involves designing questions that allow students to actively apply scientific practices. The importance of tasks that assess skills such as designing investigations, analyzing data, and arguing based on evidence is highlighted. So that their understanding of physics concepts through practical application aligns with the principles of ECD assessment (Pellegrino et al., 2016).

Based on the results of interviews with several physics teachers at one of the high schools in Pontianak City, most assessments created by teachers only evaluate each student's calculation skills without measuring deeper understanding. This is due to the lack of skills among teachers in designing assessments and their incomplete understanding of assessment concepts and correct assessment principles, resulting in assessments that are less capable of providing opportunities for students to improve their abilities. Thus, assessment development needs to be carried out to assist teachers in designing assessments that align with learning objectives, especially for the topic of static fluids.

2. METHOD

The research and development (R&D) method in this study aims to produce a product in a specific field, followed by specific ancillary products, and then tests the effectiveness of those products (Saputro, 2016). This study applies two types of approaches: the quantitative approach and the qualitative approach. In this case, we use the quantitative approach to analyze the content validity and the results of small-scale trials, which aim to measure the achievement of learning objectives through static fluid assessments. Meanwhile, we use the qualitative approach to describe the results of the assessment analysis. This research applies the ADDIE development model, which is a systematic approach to instructional design by providing a framework to create effective learning experiences through a series of iterative steps: Analysis, Design, Development, Implementation, and Evaluation (Branch, 2009). This research only reaches the "develop" stage, which produces a valid assessment for achieving learning objectives in static fluids. Figure 1 illustrates the stages of the research on developing assessments using Evidence-Centered Design (ECD) in static fluid material

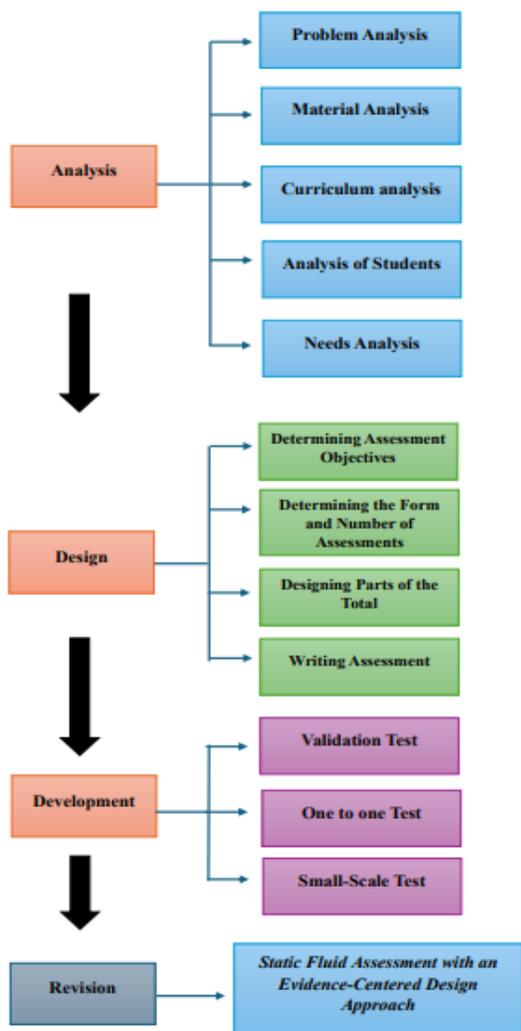


Figure 1. Flow of static fluid assessment development using ECD

The first step in this development begins with analysis, encompassing problem analysis, material analysis, curriculum analysis, and learner analysis, with the aim of designing appropriate and effective assessments. The problem analysis shows that physics assessments generally focus more on the use of formulas and calculations, without truly measuring students' deep understanding of concepts. Curriculum analysis is conducted to obtain an overview to facilitate the development of indicators that align with the Merdeka Curriculum. In addition, this analysis is expected to ensure that the assessments developed align with the learning objectives. The analysis of students is conducted

to ensure that the assessment developed meets the needs of 11th-grade high school students. Meanwhile, the content analysis establishes static fluids as the main concept in this assessment.

At the design stage, the goal is to make the assessment plan clearer and more specific. The steps include determining the assessment objectives, the format and number of questions, the assessment components, and the creation of the assessment itself. In this study, it is important to establish assessment objectives that align with learning needs and the achievement of learning goals. Scientists formulate learning objectives based on Scientific Practice (SP) competencies related to science competencies (Lager & Lavonen, 2023). The next step is to determine the type and number of assessments. In this case, we have created eight essay-style questions. Establishing the assessment components that will serve as guidelines is the first step in designing an assessment.

This assessment consists of three main components based on ECD, namely claims, evidence, and questions. Claims consist of 1) learning outcomes that align with the Merdeka Curriculum; 2) learning objectives derived from learning outcomes and linked to scientific practice (SP). SP emphasizes scientific skills, particularly the processes experienced by students in science activities (Lavery et al., 2016). Then the evidence includes: 1) Knowledge, skills, and abilities that students need to achieve learning objectives; 2) Evidence, which is the most important element in this assessment and shows that students have achieved the learning objectives. Meanwhile, the components of the questions comprise: 1) question characteristics, which dictate the design, focus, and application of the questions within a specific context; 2) questions specifically designed to measure learning objectives; and 3) scoring guidelines, which serve as guidelines for evaluating students' work based on predetermined criteria.

The final step in the design phase is to create an assessment, which is made based on the

components that have been previously established. By following these steps, it can facilitate the design of a structured assessment and provide an accurate evaluation of students' understanding of static fluid material.

At the development stage, several steps are taken, namely assessment validation, one-to-one trials, small-scale trials, and revisions based on trial results. The assessment validation process involves three lecturers or practitioners of physics education as well as three high school physics teachers. To determine the level of content validity, the Aiken formula is used, which is formulated as follows.

$$V = \frac{\sum S}{n(c - 1)} ; s = r - l_0$$

After the validation process is complete, the assessment is revised based on suggestions and feedback from the validators. Then, the assessment is tested one-to-one to obtain feedback from the students. After that, the assessment was tested on a small scale involving 10 students. The results of this trial were used to make the final revisions to the assessment.

3. RESULT AND DISCUSSION

3.1 Analysis Stage

The analysis stage is the initial phase in the development of assessments using the Evidence-Centered Design (ECD) approach in static fluids. At this stage, it includes problem analysis, material analysis, curriculum analysis, and learner analysis. The problem analysis aims to identify obstacles in the learning assessment. In this analysis, a literature review was first conducted, based on the findings of Suwandani et al. (2020), which stated that teachers experience difficulties in creating and developing assessment instruments, resulting in many assessment instruments being left unfilled and some teachers assessing students without using assessment instruments. The lack of understanding and ability of teachers in developing instruments makes them only quote what is in the guidebooks.

Arta (2024) asserts that teachers lack understanding of correct concepts and principles of assessment, leading to invalid assessments. Consequently, these assessments are less effective in measuring and supporting the achievement of learning objectives.

This stage is supported by the results of interviews with high school physics teachers in Pontianak. In the interview, the teachers revealed that the assessments they conducted were more focused on memorization and the use of formulas. As a result, the assessment is less effective in measuring the achievement of learning objectives. One of the teachers said, "I usually create assessments without considering the previously established learning objectives." This may be my mistake and oversight in designing appropriate assessments. Additionally, the lack of guidelines in designing assessments makes me more often take questions from books or previous exams.

Material analysis for the development of static fluid assessments requires a deep understanding of basic concepts such as pressure, buoyancy, density, and fluid viscosity. This assessment is designed to evaluate students' ability to apply static fluid principles to real-life situations, including solving calculations related to static fluids. Through the Scientific Practice (SP) approach, this assessment aims to measure not only students' knowledge but also their ability to apply it practically.

Curriculum analysis is conducted to ensure that assessments align with the Merdeka Curriculum, which establishes learning outcomes in the form of understanding concepts and principles of static fluids in problem-solving. Thus, this assessment is designed to align with the established learning outcomes.

The next step, student analysis, is carried out to determine students who will be the test subjects in the study. This study focuses on grade XI students as the main subjects for the development of assessments with the ECD approach. The

selection of grade XI is based on the curriculum, where the concept of static fluids is taught at this level. In addition, grade XI students are considered to have more mature knowledge and skills than previous classes, so they are more prepared to take complex assessments that are relevant to real contexts.

Finally, a needs analysis is conducted to ensure that the assessment supports learning focused on conceptual understanding, rather than mere memorization or formula usage. This analysis aims to determine the type of assessment needed, based on the results of the analysis of problems, materials, curriculum, and students. The results of this analysis indicate that the assessment created must serve as a guide for teachers in measuring the achievement of learning objectives.

3.2 Design Stage

The design stage is the second stage in this research, which aims to create a more specific assessment design. In this design phase, there are several steps, including: 1) determining the assessment objectives; 2) determining the format and number of questions; 3) designing assessment components; 4) compiling the assessment. This stage aims to ensure that the assessment not only meets quality standards but also aligns with learning needs. This systematic approach helps teachers design effective assessments to evaluate various aspects of learning, as will be explained in the next section. The first step is to determine the assessment objectives. This assessment is designed to serve as a guideline for teachers in designing effective and appropriate instruments, to be an evaluation tool for measuring students' abilities, especially in static fluid materials, and to measure the achievement of previously set learning objectives. To effectively evaluate student learning outcomes, appropriate assessments are needed so that the data and information obtained meet expectations (Safitri et al., 2017).

The second step is to determine the form and number of questions. The chosen form of

questions is essays, as this type of question can help students develop a deep understanding, analyze problems, present arguments in a structured manner, and demonstrate critical thinking and the ability to solve complex problems. (Brookhart, 2010). The number of questions is set at eight, designed to be completed within 60 minutes. This number was chosen considering the difficulty level of the questions and the students' abilities, ensuring they have enough time to read, understand, and answer each question well (Arikunto, 2012).

The third step is to design the assessment components. Static fluid assessments are designed based on the components of Evidence-Centered Design (ECD), namely claims, evidence, and questions. The components found in the claim are the physics learning outcomes for static fluid material. The learning outcomes for static fluid material are that students can apply the concepts and principles of static fluids in solving problems. By integrating scientific practices, assessments can measure not only content knowledge but also the ability to apply and communicate that knowledge in real-world contexts (Pellegrino et al., 2014). Scientific Practice (SP) includes seven main components to develop scientific competence, but the focus is directed towards using and developing representations, analyzing and interpreting data, explaining with evidence, and evaluating information. The evidence component includes the knowledge, skills, and abilities that students possess to achieve learning objectives. This evidence serves as the basis for demonstrating students' success in achieving those goals. Meanwhile, the question components are designed to assess learning achievements. This component includes types of questions, statements, and scoring guidelines designed to make the evaluation process more accurate and objective.

The final step in this process is to prepare the assessment. In this process, the material or topic becomes an important element in measuring students' abilities in static fluids. The topics used in preparing the assessment can be seen in Table 1.

Table 1. Static fluid assessment topic

Topic	Question Number
Hidrostatic Pressure	5 and 7
Archimedes Principle	1, 2 and 6
Pascal's Law	4 and 8
Viscosity	3

Each part of the assessment is designed to evaluate the extent to which students understand the concepts that have been taught. Thus, the assessment can provide a clear picture of the expected competency achievements. Examples of the assessment component arrangement can be seen in Table 2.

Table 2. Example of static fluid assessment components using ECD

Learning Outcomes	Student can apply the concepts and principles of static fluids in solving problems.
Scientific Practice	SP4 (Analyzing and interpreting data)
Learning Objectives	Students can analyze the data on the relationship between F (Force) and A (cross-sectional area) to identify which category can lift an object with a large cross-section.
Knowledge, Skills, and Abilities	<ol style="list-style-type: none"> 1. Able to identify variables in the formula 2. Able to apply basic physics concepts in real-life situations 3. Able to read and analyze data correctly 4. Writing Pascal's law formula
Evidence	<ol style="list-style-type: none"> 1. Calculating to magnitude of the pressure ratio applied 2. Conclude which category can lift objects on a large cross-section
Task Framework	Conclude which category can lift objects on a large cross-section

Assessment components designed to measure question characteristics play an important role in evaluating students' understanding of the material that has been taught. In this assessment, questions must be designed in such a way that they reflect the competencies to be measured, such as the ability to analyze and interpret data. Therefore, it is important to ensure that each type of question aligns with the relevant assessment

characteristics so that the evaluation objectives can be effectively achieved. The questions created must be able to accurately measure the students' abilities. A useful question not only tests students' understanding but also encourages critical thinking skills in problem-solving. As for examples of questions to assess students' ability to analyze and interpret data, they can be seen in Figure 2.

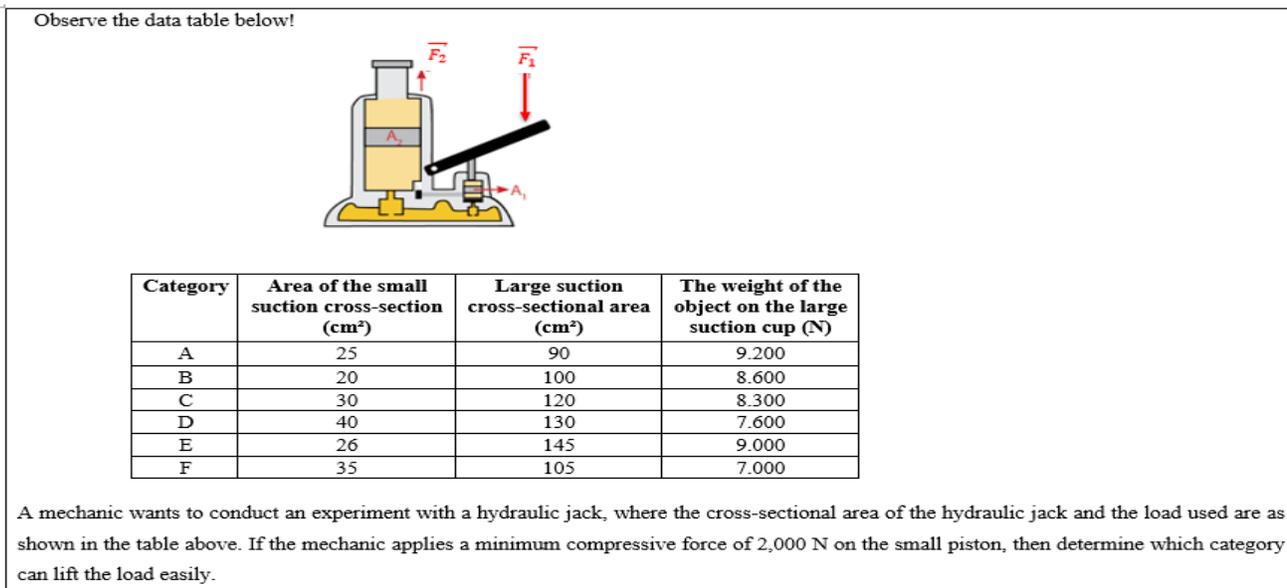


Figure 2. Example questions for analyzing and interpreting data

3.3 Development Stage

This stage represents the third phase of the assessment validation process. The steps taken at this stage include: 1) conducting an assessment validation test; 2) conducting a one-to-one test; and 3) conducting a small-scale trial. This stage aims to validate the created instrument and serve as a guide for teachers in crafting suitable assessments.

The first step, assessment validation test, is carried out to measure whether the assessment is appropriate and valid and meets the criteria. (Arikunto Suharismi, 2014) asserts that the validation stage strives to verify the research instrument's compliance with validity criteria, evaluating it both logically and empirically, before approving its use in research. Content validation is carried out in two stages. Three panel validators, consisting of high school physics teachers, assess the content in the first stage, while three expert validators from among physics education lecturers conduct the second stage. This validation process is necessary so that the instrument can function as an effective and objective measuring tool (Sirait et al., 2023).

Based on the validation results, the validator offered numerous inputs and recommendations for the developed assessment. In the validation process, the high school physics teacher validator provided feedback to use sentences that adhere to Indonesian language rules, correct typos, clarify the use of formulas for each concept in the answer key, and ensure that the question instructions are clear. Additionally, the physics education lecturer validator suggested improving the use of vector symbols to be more consistent with other questions and ensuring that the use of images, sketches, graphs, and tables is clearer and easier for students to understand. In this case, it can enhance the way questions are presented and facilitate students' understanding of the provided information.

In the validation process, interviews were conducted with the three physics teacher validators in the development of this assessment. Their answers about this test show that the new thing about the ECD instrument is that it looks at the evidence that students are supposed to give. This means that the testing is more focused on what students can show correctly and in line with the skills, knowledge, and abilities that come from

the learning goals. During the interview, the teachers expressed their appreciation for the ECD approach assessment, stating, "I believe this approach is highly relevant to curriculum development, as it allows students to freely express their understanding through open-ended question types and requires them to present their answers in various formats." They went on to say, "In my opinion, the ECD approach assessment has the potential to demonstrate that students have met the established learning objectives." Students' critical thinking skills are indirectly required when they work on this ECD-based test instrument, as the presence of evidence components in assessments using the ECD approach serves as indicators of students' achievement of learning objectives.

In addition to receiving suggestions for improvement, the validator also evaluates the

content of the assessment. The results obtained from the validators were then analyzed using the Aiken formula on the aspects of content, construction, and language. The calculation results for the content aspect yielded an Aiken index with an average of 0.88, which was classified as very good. These results demonstrate the high suitability of the assessment's content for the students' needs and learning objectives. Then, for the construction aspect, an average score of 0.87 was obtained, categorized as very good. This result suggests a clearly structured assessment. Next, the language aspect yielded an average Aiken index score of 0.86, qualifying it as very good. This result demonstrates the use of appropriate and comprehensible language. The overall Aiken index, which measures content, construction, and language, averages 0.87, indicating very good quality. Table 3 displays the content validation scores for each question.

Table 3. Assessment validation results

Scientific Practice (SP)	Question Number	Value	Category	Explanation
Developing and Using Representation	1	0,87	Very High	Valid
	2	0,86	Very High	Valid
Analyzing and Interpreting Data	3	0,90	Very High	Valid
	4	0,88	Very High	Valid
Explaining with Evidence	5	0,87	Very High	Valid
	6	0,87	Very High	Valid
Evaluating Information	7	0,85	Very High	Valid
	8	0,88	Very High	Valid

Based on the overall results, the assessment is declared valid and falls into the "Very High" category. The high content validity indicates that the assessment is overall relevant to the learning objectives it aims to measure (Sugiyono, 2013). With very high validity, this assessment is declared valid and can be used to measure each student's ability, especially in understanding static fluid concepts (Yusup, 2018).

After undergoing a series of validations and revisions, the revised assessment was then tested in a one-to-one trial with one student as a sample over a period of 60 minutes. The purpose of the one-to-one trial is to evaluate student responses

and uncover technical errors, ambiguities, or inconsistencies in the instrument before it is widely used (Reiser et al., 2024). According to Dick et al. (2015), one-to-one trials are essential to ensure that participants clearly understand the instructions, questions, or concepts. The one-to-one trial showed that students understand most questions, but deeper analysis questions are difficult. This is due to the students' still insufficient understanding of static fluid material. In terms of the time given, students feel that 60 minutes is sufficient to complete the assessment. The results of the one-to-one trials and student feedback indicate that the assessment created is deemed suitable for a small-scale trial involving

10 eleventh-grade students. This small-scale trial aims to re-evaluate the effectiveness of the questions and ensure that the completion time is more representative (Sirait & Oktavianty, 2021). After the trial, students were asked to fill out a questionnaire to provide feedback on the assessment they had completed. The survey results show that all students found the instructions and questions provided to be very clear, and they did not experience any difficulties in answering them.

The results of the questionnaire after a small-scale trial showed that 8 out of 10 students felt that 60 minutes was sufficient to complete all the questions thoroughly. However, two students felt that the time was not enough because some questions were considered difficult and required more time to understand. This indicates a difference in time requirements among students, which is likely influenced by the difficulty level of the questions, their understanding, and each student's speed in completing them. Based on these results, the assessment is considered ready to be applied to a larger group, with the note that the questions must be ensured to be clear and the working time sufficient for all students.

Following a small-scale trial, students provided clear answers accompanied by supporting evidence. The Evidence-Centered Design (ECD) approach-developed assessments encourage students to focus on the presentation of evidence, leading to more purposeful task completion. These findings indicate that the ECD approach makes assessments more effective in uncovering students' abilities and conceptual understanding. Thus, it is expected to serve as a reference for teachers in designing appropriate and effective assessments. The assessment developed aims to assist teachers in conducting evaluations and serves as a clear and structured guideline in assessing each student's understanding and abilities. Subsequently, this assessment can serve as a reference for teachers in designing appropriate and effective assessments. Thus, teachers can more easily assess students' learning

outcomes and develop appropriate learning strategies.

4. CONCLUSION

Based on the findings of research and development, it was concluded that this study has successfully developed an assessment using the Evidence-Centered Design (ECD) approach, which is suitable for improving the quality of physics learning assessments. Three stages determine the feasibility category in this case: 1) Six experts validate the assessment, resulting in an Aiken index of 0.87 and a very high (valid) classification; 2) One-to-one testing reveals clear questions and ample time; and 3) Results from small-scale tests and questionnaires demonstrate the appropriateness and clarity of the ECD approach's questions, along with ample time. ECD-based assessments can serve as a guideline for teachers in creating more structured, effective assessments that are relevant to the needs of 21st-century learning.

5. REFERENCES

- Arieli-Attali, M., Ward, S., Thomas, J., Deonovic, B., & von Davier, A. A. (2019). The Expanded Evidence-Centered Design (e-ECD) for Learning and Assessment Systems: A Framework for Incorporating Learning Goals and Processes Within Assessment Design. *Frontiers in Psychology, 10*. <https://doi.org/10.3389/fpsyg.2019.00853>
- Arikunto, S. (2012). *Dasar-dasar Evaluasi Pendidikan* (3rd ed.). Bumi Aksara.
- Arikunto Suharismi. (2014). *Prosedur Penelitian Suatu Pendekatan Praktik*. Rineka Cipta.
- Arta, G. Y. (2024). Asesmen dalam Pendidikan: Konsep, Pendekatan, Prinsip, Jenis, dan Fungsi. *Jurnal Pendidikan, Bahasa Dan Budaya*.
- Bao, L., & Koenig, K. (2019). Physics education research for 21st century learning. *Disciplinary and Interdisciplinary Science Education Research, 1*(1), 1–12.

- <https://doi.org/10.1186/s43031-019-0007-8>
- Bechard, S., Clark, A., Swinburne Romine, R., Karvonen, M., Kingston, N., & Erickson, K. (2019). Use of Evidence-Centered Design to Develop Learning Maps-Based Assessments. *International Journal of Testing*, 19(2), 188–205. <https://doi.org/10.1080/15305058.2018.1543310>
- Branch, R. M. (2009). *Instructional Design: The ADDIE Approach*. Springer, United States of America. Springer.
- Brookhart, S. M. (2010). *How to Assess Higher-order Thinking skills in Your Classroom*. ACSD Member Book.
- Dick, W., Carey, L., & Carey, O. J. (2015). *The Systematic Design of Instruction* (8th ed.). Pearson.
- Duschl, R. A., & Bybee, R. W. (2014). Planning and carrying out investigations: an entry to learning and to teacher professional development around NGSS science and engineering practices. *International Journal of STEM Education*, 1(1), 1–9. <https://doi.org/10.1186/s40594-014-0012-6>
- Indriani, N., Sutamah, S., Maharani, A. P., Wahyuni, N., Afi, D. A., Bahri, F. F., & Rahma, S. N. (2023). Analisis Kesulitan Guru Dalam Mengembangkan Instrumen Asesmen Autentik Higher Order Thinking Skill Di MI Sidorajo. *Afeksi: Jurnal Penelitian Dan Evaluasi Pendidikan*, 4(6), 596–607. <https://doi.org/10.35672/afeksi.v4i6.171>
- Kellaghan, T., & Greaney, V. (2001). *Using Assessment To Improve the Quality of Education*. UNESCO International Institute for Educational Planning.
- Kistiono. (2019). Pengembangan Tes Kemampuan Berpikir Tingkat Tinggi Fisika SMA. *Jurnal Inovasi Dan Pembelajaran Fisika*, 6(1).
- Kurniawati, A., & Sukardiyono, S. (2018). The Development of Authentic Assessment Instrument to Measure Science Process Skill and Achievement based on Students' Performance. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, 4(2), 65–74. <https://doi.org/10.21009/1.04203>
- Lager, A., & Lavonen, J. (2023). Engaging Students in Scientific Practices in a Remote Setting. *Education Sciences*, 13(5), 431. <https://doi.org/10.3390/educsci13050431>
- Lane, S. (2015). *Handbook of Test Development*. Routledge. <https://doi.org/10.4324/9780203102961>
- Laverty, J. T., Underwood, S. M., Matz, R. L., Posey, L. A., Carmel, J. H., Caballero, M. D., Fata-Hartley, C. L., Ebert-May, D., Jardeleza, S. E., & Cooper, M. M. (2016). Characterizing College Science Assessments: The Three-Dimensional Learning Assessment Protocol. *PLOS ONE*, 11(9), e0162333. <https://doi.org/10.1371/journal.pone.0162333>
- Maulani, N., Linuwih, S., & Sulhadi, S. (2020). Analisis kemampuan pemecahan masalah fisika dalam asesmen higher order thinking. *Prosiding Seminar Nasional Pascasarjana*, 590–597.
- Mislevy, R. J., & Haertel, G. D. (2006). Implications of Evidence-Centered Design for Educational Testing. *Educational Measurement: Issues and Practice*, 25(4), 6–20. <https://doi.org/10.1111/j.1745-3992.2006.00075.x>
- Mislevy, R. J., & Riconscente, M. M. (2005). *Evidence-Centered Assessment Design: Layers, Structures, and Terminology*. SRI International.
- Newton, S., Alemdar, M., Rutstein, D., Edwards, D., Helms, M., Hernandez, D., & Usselman, M. (2021). Utilizing Evidence-Centered Design to Develop Assessments: A High School Introductory Computer Science Course. *Frontiers in Education*, 6. <https://doi.org/10.3389/feduc.2021.695376>
- Nur'aini, D. A., Lestari, P. D., & Kurniawan, B. R. (2020). Pengembangan Asesmen Formatif Berbasis Komputer Untuk

- Mengetahui Penguasaan Konsep Siswa Pada Materi Hukum Bernoulli. *Pengembangan Asesmen Formatif Berbasis Komputer Untuk Mengetahui Penguasaan Konsep Siswa Pada Materi Hukum Bernoulli*, 5(2), 106–112.
- Pellegrino, J. W., DiBello, L. V., & Goldman, S. R. (2016). A Framework for Conceptualizing and Evaluating the Validity of Instructionally Relevant Assessments. *Educational Psychologist*, 51(1), 59–81. <https://doi.org/10.1080/00461520.2016.1145550>
- Pellegrino, J. W., Wilson, M. R., Koenig, J. A., & Beatty, A. S. (2014). *Developing assessments for the next generation science standards*. National Research Council.
- Reiser, R. A., Carr-Chellman, A. A., & Dempsey, J. V. (2024). *Trends and Issues in Instructional Design and Technology*. Routledge. <https://doi.org/10.4324/9781003502302>
- Safitri, A. N., Sari, R., & Wahyuni, S. (2017). The Influences of Mathematics Ability toward Physics Learning in Senior High School Based on an Authentic Assessment System. *International Journal of Learning and Teaching*, 3(1), 11–14. <https://doi.org/10.18178/ijlt.3.1.11-14>
- Salassa, S., Amin, B. D., & Ali, M. S. (2023). Kesulitan Materi Fisika Pada Peserta Didik Kelas X MIPA SMA Negeri 14 Bulukumba. *Jurnal Sains Dan Pendidikan Fisika*, 19(1), 69–88. <https://doi.org/10.35580/jspf.v19i1.32349>
- Saputro, B. (2016). *Manajemen Penelitian Pengembangan (Research & Development) Bagi agi Penyusun Tesis dan Disertasi*. Aswaja Pressindo.
- Sirait, J., Firdaus, F., Hidayatullah, M. M. S., & Habellia, R. C. (2023). Development and Validation of Force Test to Assess Physics Education Students' Representational Competence. *Jurnal Pendidikan Sains Indonesia*, 11(2), 306–317. <https://doi.org/10.24815/jpsi.v11i2.28178>
- Sirait, J., & Oktaviany, E. (2021). Pengembangan dan Validasi Angket Motivasi Belajar Fisika (AMBF): Studi Pilot. *Jurnal Ilmiah Pendidikan Fisika*, 5(3), 305. <https://doi.org/10.20527/jipf.v5i3.3829>
- Sugiyono, S. (2013). *Metode Penelitian Kualitatif Kualitatif dan R&D*. Alfabeta.
- Sukmawa, O., Rosidin, U., & Sesunan, F. (2019). Pengembangan Instrumen Asesmen Kinerja (Performsnce Assessment) Praktikum Pada Mata Pelajaran Fisika Di SMA. *Jurnal Pendidikan Fisika*, 7(1), 116. <https://doi.org/10.24127/jpf.v7i1.1397>
- Suwandani, R. A., Karma, I. N., & Affandi, L. H. (2020). Analisis Faktor-Faktor Kesulitan Guru Dalam Melaksanakan Penilaian Autentik di SDN Gugus 1 Kecamatan Janapria. *Jurnal Ilmiah Pendidikan Indonesia*, 2(1).
- Yusup, F. (2018). Uji Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif. *Jurnal Tarbiyah: Jurnal Ilmiah Kependidikan*, 7(1), 17–23. <https://doi.org/10.18592/tarbiyah.v7i1.2100>
- Yuwarti, Y., Pasaribu, M., & Hatibe, A. (2017). Analisis Pemahaman Konsep Vektor pada Siswa SMA Negeri 5 Palu. *JPFT (Jurnal Pendidikan Fisika Tadulako Online)*, 5(3), 12. <https://doi.org/10.22487/j25805924.2017.v5.i3.8865>
- Zieky, M. J. (2014). An introduction to the use of evidence-centered design in test development. *Psicología Educativa*, 20(2), 79–87. <https://doi.org/10.1016/j.pse.2014.11.003>