QUANTUM: Jurnal Inovasi Pendidikan Sains, Vol. 16, No. 2, 257-267
p-ISSN : 2086-7328 e-ISSN : 2550-0716
Website: https://ppjp.ulm.ac.id/journal/index.php/quantum/index
DOI: 10.20527/quantum.v16i2.21163

DEVELOPMENT OF ASSESSMENT USING *EVIDENCE-CENTERED DESIGN* (ECD) ON DYNAMIC FLUIDS MATERIAL

Sidik Wijaya, Judyanto Sirait, Ray Cinthya Habellia
Physics Education Study Program, Faculty of Teacher Training and Education, Tanjungpura University
Jl. Prof. Dr. H. Hadari Nawawi, Pontianak 78115, Kalimantan Barat, Indonesia
*corresponding author: f1051211046@student.untan.ac.id

Abstract. This research aims to develop an assessment using the Evidence-Centered Design (ECD) approach for dynamic fluid materials, employing the Research and Development (R&D) research method. This assessment research and development uses the ADDIE model, which was limited to the development stage to produce an assessment. The content validity test result for the questions, using Aiken's value, was 0.834, which is categorized as very high and valid. The data collection process in assessment development was carried out through validator input, assessment trial implementation, and student responses obtained through questionnaires as a basis for revision to produce a valid, reliable, and relevant instrument. Data were obtained through a limited trial with one student and a small-scale trial with one class of 12th-grade students using questionnaires. The results of the students' responses indicate that the Evidence-Centered Design (ECD)-based assessment is considered clear, relevant, and helpful for understanding concepts, although some questions were considered difficult. From this series of processes, a dynamic fluid assessment was produced as one alternative for teachers to measure student competency and develop evidence-based assessments.

Keyword: assessment, evidence-centered design (ECD)

INTRODUCTION

Physics is a subject in high school within the science group that is empirical, as it is based on the observation of natural phenomena (Hanum, 2021). One of the branches studied is fluid dynamics, which includes ideal fluids, the continuity equation, and Bernoulli's equation, along with their applications in daily life (Affandy et al., 2019). In physics learning, it is important to implement a deep and comprehensive assessment system, one of which is the design of an assessment. Assessment in the learning process in the current era should not be viewed solely as a stage for generating student grade reports. However, assessment should be understood as an activity to collect and process information in order to determine the needs, learning progress, and learning outcomes of students (Damayanti & Untari, 2024). Therefore, the next discussion will focus on the discussion of assessment.

Assessment, measurement, research, and evaluation are interconnected in the scientific process. Assessment serves to gather data, measurement quantifies it, research utilizes it to understand phenomena, and evaluation compares it against standards (Huitt et al., 2001). Assessment should not be just a formality, but must be done well so that learning outcomes and the learning process can be evaluated (Boy, 2022). Assessment is an important part of education because it is used to make decisions regarding students, curriculum, and policies (Munaroh, 2024), and it includes activities such as observing, recording, scoring, and interpreting student learning (Maemonah, 2018).

The conventional assessment model that measures achievement through multiple-choice, true/false, or matching tests is considered less effective in

Published by: Program Studi Pendidikan Kimia FKIP Universitas Lambung Mangkurat
How to Cite: Wijaya, S., Sirait, J. , & Habellia, R. C. (2025). Development of Assessment Using Evidence-Centered Design (ECD) on Dynamic Fluids Material. *QUANTUM: Jurnal Inovasi Pendidikan Sains*, *16*(2), 257–267. https://dx.doi.org/10.20527/quantum.v16i2.21163.

Article History:
Received : 13-12-2024, *Accepted* : 10-10-2025, *Published* : 22-10-2025

comprehensively representing student performance. This type of objective test is also difficult to measure aspects of attitudes, skills, and knowledge which relevant to real life (Voogt & Roblin, 2010). Assessment practices are still dominated by traditional (summative) evaluations, even though most academics emphasize the importance of formative assessment and feedback for improving learning strategies (Monteiro et al., 2021). Many teachers still face challenges in understanding learning outcomes, making it difficult to derive objectives and develop teaching modules, and the assessments they create are often not aligned with the learning objectives. This makes it difficult to map students' abilities. Meanwhile, it states that conventional learning and assessment tend to be monotonous, less engaging, and require more time and effort (Sirait et al., 2024; Septyana et al., 2024). Therefore, it is important for an educator to shift the learning assessment paradigm to be more learning-oriented, where assessment is not merely a tool for measuring, but also a tool for improving and enhancing learning.

According to Puteri *et al*. (2023), meaningful assessment should be conducted authentically, meaning students are able to demonstrate their understanding in the context of real life, and is more effective when it is formative by providing continuous feedback for both teachers and students. However, Sarwi *et al*. (2019) found that science assessment instruments are still lacking focus and do not provide appropriate information. Therefore, this study develops an Evidence-Centered Design (ECD)-based assessment that offers a structured framework from learning objectives and task design to results interpretation. Arieli-Attali *et al*. (2019) also emphasize that ECD is a comprehensive framework for integrating learning and assessment, and for supporting the development of more meaningful modern assessment systems.

Evidence-Centered Design (ECD) is a principle-based framework for designing, developing, and delivering educational assessments. The main focus is on assessments that utilize complex student models and interactive simulations. The main contribution of ECD lies in providing a coherent framework for integrating various assessment developments toward a common goal. (Mislevy *et al*, 1999). In line with this approach, it is necessary to have *Scientific Practices*, which is a framework that expects performance to combine relevant science or engineering practices with core disciplinary ideas and crosscutting concepts appropriate for students at a specified grade level (NGSS Lead States, 2013). Therefore, the ECD framework can be seen as an effort to integrate the assessment process with scientific practices, enabling the design of more innovative assessments that are aligned with students' cognitive knowledge needs.

This research aims to develop an assessment that utilizes the *Evidence-Centered Design* (ECD) framework integrated with *scientific practice*. By combining the principle-based and evidence-based approach of ECD with integration into *scientific practice*, the resulting assessment is expected not only to accurately measure students' abilities but also to support a deeper, more reflective, and concept- and evidence-based learning process. This development will contribute to the creation of assessments that are more relevant and aligned with the challenges of contemporary education, providing a comprehensive understanding of students' cognition in a more authentic context.

METHOD

This research uses the *Research and Development* (R&D) approach to develop an *Evidence-Centered Design (*ECD)-based assessment product for dynamic fluid material. The R&D method was chosen because it focuses on developing new

products and testing their effectiveness (Sugiyono, 2016). The R&D approach was selected because it aligns with the research objectives, which focus on developing and validating a new product.

Product development and educational research require a systematic framework to ensure the desired goals are achieved. Various development models have been introduced, one of which is the frequently used ADDIE model. The ADDIE model, which consists of five main stages: a*nalysis, design, development, implementation,* and *evaluation*, provides a structured process for designing and developing educational programs or products (Amaliah *et al*., 2023). However, the limitations of this study only extend to the *Development* stage to produce an ECD-based assessment prototype. The stages in this ADDIE method are outlined in the following table:

Table 1. Stages of the research procedure using the ADDIE model

No	Framework Concept	Procedure
1	**Analyze** Identify the causes of learning problems in pre-planning, thinking about and deciding on the lessons and/or subjects or courses to be taught (Branch, 2009).	1. Analysis in identifying learning needs. 2. Analysis of learning materials.
2	**Design** Verify the desired results or achievements (learning objectives) and determine the methods or strategies to be applied (Branch, 2009).	1. Formulating assessment objectives. 2. Determining the structure of the assessment framework. 3. Designing the assessment framework using Evidence-Centered Design (ECD). 4. Visual design and technical arrangements. 5. Initial assessment design framework.
3	**Develop** Develop and validate learning resources, as well as develop the necessary supporting materials and strategies (Branch, 2009).	1. Creating assessment questions. 2. Assessment validation. 3. First-stage revision. 4. Limited trial (small-scale tryout). 5. Second-stage revision. 6. Small-scale trial. 7. Final revision.

Analyze

The *analysis* stage is a crucial step in developing *Evidence-Centered Design* (ECD)-based assessments because it lays the foundation for the entire development process. The assessment development process begins with an initial needs analysis to establish competencies, test-takers, and test objectives (Nawawi & Wijayanti, 2018).

A structured initial analysis was conducted to identify any problems or obstacles that arose related to learning (Hajati & Amaliah, 2022). At this stage, several steps of the analysis phase are carried out, including: Task analysis is performed to determine the necessary tasks to achieve learning outcomes (Norlaila *et al*., 2024), learning needs analysis begins with an examination of learning needs and problems, curriculum content analysis, and analysis of high school/Islamic high school textbook content (Zural *et al*., 2024), analyzing problems that occur during learning, analyzing student characteristics and needs, analyzing the characteristics of professional education course material, and mapping the relationship between material substance and assessment instruments (Amaliah *et al*., 2023).

Design

The *designing* stage is a critical stage where abstract concepts from the analysis are transformed into concrete plans that will be used in the development of the assessment. At this stage, the researcher focuses on designing the structure and essential elements of the assessment, integrating the principles of *Evidence-Centered Design* (ECD) to ensure the assessment is valid, reliable, and capable of providing meaningful information about student learning.

In the process of developing assessment instruments, the design stage plays a crucial role. This stage is a follow-up step after the analysis conducted in the previous stage, which is the analysis stage. According to Rahma (2012), "The ability and creativity of educators in designing assessment instruments play an important role in optimizing the achievement of learning indicators." Therefore, the second stage focused on designing assessment instruments tailored to the characteristics of the material.

Develop

The *development* stage is the core of *the Evidence-Centered Design* (ECD) assessment development process for dynamic fluid material. This stage involves translating design concepts into a concrete product, namely an assessment ready for testing. The goal of this stage is to produce assessments that are valid, reliable, and ready for implementation.

At this stage, the previously designed assessment is validated by the validator. The content validity test of the questions was conducted by six expert panels, consisting of three physics lecturers and three experienced physics teachers. The purpose of this test is to ensure that the content of the developed questions is valid and relevant. This validation process is carried out by providing a validation sheet to experts. Data validity was obtained through a questionnaire using a Likert scale and calculated using Aiken's V formula.

The formula proposed by Aiken is as follows:

$$V = \frac{\sum s}{[n(C-1)]} ; s = r - Lo$$

The Aiken's V formula is used to calculate the content validity coefficient, which is based on the assessment of n expert panelists regarding an item, considering the extent to which the item represents the construct being measured (Hendryadi, 2017).

RESEARCH AND DISCUSSION

Analyisis Stage

In this stage, the researcher conducted various analysis activities to understand the needs, characteristics, and context that would influence the design and development of the assessment. Starting by identifying the learning outcomes that are intended to be achieved through assessment. In the context of dynamic fluid material, the learning achievement at the end of phase F is that students are able to apply the concept of dynamic fluids in solving problems and apply its principles.

Analyzing learning competencies including conceptual knowledge, analytical skills, and the ability to apply theory in real-world situations. Researchers collected relevant literature information to ensure that these learning outcome expectations are relevant and appropriate for the students' educational level. Starting with identifying the learning standards and objectives that have been established in the curriculum.

This step ensures that the analyzed competencies align with educational needs and are relevant to the demands of the working world. Next, a competency structure analysis was conducted by breaking it down into the necessary skills, knowledge, and attitudes.

The analysis of learning materials was conducted through concept mapping to identify the main principles of fluid dynamics, including flow rate, the principle of continuity, and Bernoulli's law. This mapping is presented in the form of a concept diagram to organize the relationships between concepts and identify areas requiring special emphasis in assessment. Additionally, the analysis also includes determining the complexity level of the material, considering that the topic of fluid dynamics requires strong mathematical abilities and spatial visualization. The parts of the material that were considered challenging for the students were then identified so that the assessment could be designed to be targeted appropriately. Next, real-life examples of dynamic fluid applications are integrated into the assessment to enhance the relevance, engagement, and conceptual understanding of the learners. Morrison (2001: 27) divides the function of needs analysis by identifying needs relevant to the current job or task, namely what problems affect learning outcomes (Abidin, 2007).

The purpose of the above analysis stage is to align the designed assessment with the learning outcomes. Learning achievement at the end of Phase F is that students are able to apply vector concepts and principles to the kinematics and dynamics of particle motion, work and energy, fluid dynamics, simple harmonic motion, sound waves, and light waves in solving problems, as well as applying the principles and concepts of heat energy and thermodynamics with its various changes in heat engines (Kemdikbud, 2022).

Designing Stage

In the design phase, the researcher focuses on designing the structure and essential elements of the assessment, integrating ECD principles to ensure the assessment is valid, reliable, and capable of providing meaningful information about student learning. This stage involves several important steps, including: formulating assessment objectives, determining the structure of the assessment framework, designing the assessment framework using *Evidence-Centered Design* (ECD), visual design and technical setup, and creating an initial assessment prototype.

Based on the results of the needs analysis, the researcher identified the learning objectives formulated to serve as a reference in the development of assessments. The assessment reference is based on the improvement of students' abilities in two parts: understanding of dynamic fluid materials and process skills (observing, questioning and predicting, planning and conducting investigations, processing and analyzing data, creating, evaluating, and communicating results). This reference is based on the learning achievements of students in Phase F.

Based on Scientific Practices, which is a framework that expects performance to combine relevant science or engineering practices with core disciplinary ideas and crosscutting concepts appropriate for students at a specified grade level (NGSS Lead States, 2013). At this stage, the researcher selects the most appropriate question format to achieve the assessment objectives, which is to reflect how students understand and apply the concepts of fluid dynamics. The researchers then determined the overall structure of the assessment. This structure is designed to cover a range of difficulty levels and task types, providing a comprehensive overview of student understanding. Additionally, the researchers designed clear and easy-to-understand instructions for each type of question, providing sufficient guidance so that students could answer the questions correctly without confusion.

The design of the *Evidence-Centered Design* (ECD)-based assessment framework is carried out through several stages. First, establishing the knowledge and skills domains to be measured, namely students' understanding of flow rate, continuity, and Bernoulli's law. Second, determining assessment requirements including the framework structure, task models, and scoring rubrics. Third, developing an assessment framework based on the main components of ECD, namely the *student model* which describes the competencies being measured, the *evidence model* which identifies evidence to assess these competencies, and the *task model* which designs contextual tasks or questions to represent the application of knowledge in real-world situations.

Developing the visual presentation of assessments, including page layout, fonts, and graphic elements that aid in understanding the questions. The visual design not only should be appealing but also functional, ensuring learners can focus on the content without distraction. After all the above elements have been designed, the next step is to develop the initial assessment design framework. This initial assessment design framework is a physical representation of all the planning that has been done and is used for further evaluation before the trial run. One of the initial assessment design frameworks that has been designed is as follows:

Table 2. Initial assessment design framework

Learning Achievement	Students are able to apply the concepts and principles of fluid dynamics to solve problems.
Scientific Practices 4th	Analyze and interpret data
Learning Objectives	Students can correctly apply the principle of continuity to fluid flow and understand the volumetric flow rate of fluids in closed pipes.
Knowledge, Skills, and Abilities	• Know the principle of continuity in fluid flow. • Able to write the continuity equation. • Able to calculate the flow velocity of water in a pipe.
Evidence	• Students calculate the flow velocity of water in pipes of different sizes using the continuity equation. • Students explain the consequences of changes in water velocity in a pipe.
Characteristics of Questions	A discourse illustration, an image illustration, and data in the form of a table are presented. Asking students to calculate the flow rate of water in other pipes and to explain the changes in fluid velocity and its effect on other pipes with accurate calculations.

Development Stage

A more detailed explanation of the steps involved in the Development stage is as follows: creating assessment questions, validating the assessment, first-stage revision, small-scale tryout, second-stage revision, small-scale testing, and final revision.

Based on the initial assessment design framework that was created in the design phase, the first step in development is creating the assessment questions. This assessment covers various types of questions designed to measure different aspects of dynamic fluid concepts. The type of question used is an essay. Each question is designed using the *Evidence-Centered Design* (ECD) approach, where each question item is linked to a specific evidence model. Here are example questions from an assessment designed using ECD:

A city is designing a PDAM (*local water supply utility*) pipeline system to distribute clean water to residents' homes. Air is flowed from the central tank through pipes of different sizes. The main PDAM pipe is near the larger tank, while the pipes leading to residents' homes are smaller. Engineers must calculate the water flow rate and ensure sufficient water pressure so that water can reach all parts of the house, including those far from the tank.

The ilustration of PDAM pipe and residents' pipe connection

Pipe	Pipe Diameter (m)	Water Flow Rate (m/s)
PDAM Pipe	**0.14**	**2**
Residents' Pipe	**0.07**	**?**

a) To assist the engineers, your task is to calculate the water velocity in the residents' pipes.
b) B. What happens if the water pressure in the main PDAM pipe decreases, causing the water flow velocity to become 1 m/s? Explain the impact on the water flow velocity in the residents' pipes with accurate calculations.

Figure 1. Assessment question using Evidence-Centered Design

At this stage, the previously designed assessment is validated by the validator. The content validity test of the questions was conducted by six expert panels, consisting of three physics lecturers and three experienced physics teachers. The purpose of this test is to ensure that the content of the developed questions is valid and relevant. This validation process is carried out by providing a validation sheet to experts. Data validity was obtained through a questionnaire using a Likert scale. Here are the results of the assessment question validation:

Table 3. Data from the validation of dynamic fluid assessment questions

Item No	Validation Score Average	Category	Description
1	0,805	Very High	Valid
2	0,827	Very High	Valid
3	0,855	Very High	Valid
4	0,838	Very High	Valid
5	0,860	Very High	Valid
6	0,838	Very High	Valid
7	0,855	Very High	Valid
8	0,794	High	Valid

According to Aiken, the validity of an item can be considered good if the Aiken index obtained is greater than or equal to 0.75, involving eight raters and four answer categories (Retnawati in Nabil *et al*., 2022). In general, the results of the assessment of the validation sheets and research instruments indicate that the instruments are "suitable for use" and highly valid, as the suggestions and recommendations from the validators have been implemented by the researcher (Soeyono, 2014). Based on the validation results obtained, it is known that the overall average validation score for the questions is an average Aiken's value of 0.834, which is categorized as very high and valid.

After validation, the next step is to make improvements or revisions. The results of these improvements are then tested on students. (Kasdriyanto *et al.*, 2021). Assessment can be considered valid if the results of the validation assessment meet the valid category according to the established criteria. If the results are not yet valid, revisions will be made based on the suggestions and input from the validator, with

the aim of correcting the weaknesses present in the assessment (Kurnia, 2019). The first revision stage after assessment validation is an important step in ensuring the quality and reliability of the assessment. At this stage, feedback from the validation process, including expert evaluation, is used to improve the content and test items of the assessment that do not meet the established standards. The validation suggestions include improving the sentences and wording in the assessment and improving the learning objectives in the assessment.

Limited trials were conducted on one 12th-grade student. At this stage, a 12th-grade student was selected as a representative to participate in the assessment, which aimed to serve as an initial sample for testing the assessment questions. The question needs to be tested first and the results analyzed to meet the requirements of a good test (Kadir, 2015). The pilot testing phase *in Evidence-Centered Design* (ECD)-based assessment aims to evaluate and validate the instrument before it is used widely. Some statements from the questionnaire given to students are: the questions are presented in easy-to-understand language, the instructions and information in the questions are clear and not confusing, the questions have a difficulty level appropriate for my abilities, the questions give me the opportunity to think critically and creatively, the questions help me understand the concepts in the material taught, the questions are relevant to real-life situations or the application of learned concepts, the illustrations or images in the questions help me understand the questions asked, and the questions are presented with a neat and attractive appearance. Students can provide answers to the statements in the questionnaire by marking the options in a table containing the choices *strongly agree, agree, disagree, strongly disagree*. The results obtained from this trial are to see the students' responses and suggestions when working on assessment questions through the questionnaires provided. Student responses indicated that "*questions in the form of illustrations and images can help in understanding physics concepts in real life, making them interesting and creative, but overall the questions are somewhat difficult to solve*."

After limited testing on one 12th-grade student, the next step in developing *Evidence-Centered Design* (ECD)-based assessments is to conduct a second revision. This revision is based on the evaluation results from the trial, with the goal of improving the assessment instrument to make it more valid and reliable before testing it on a larger group.

A small-scale test involving 6 twelfth-grade students in the development of *Evidence-Centered Design* (ECD)-based assessments was conducted. This aims to identify potential problems and the effectiveness of the assessment before it is widely implemented. In this test, each student will be asked to complete tasks or questions designed according to ECD, which includes competency definitions, expected evidence, and the tasks students must complete. During the small-scale test, there were several student responses from the questionnaire provided, namely that some questions were easy to understand and solve, while others were difficult, the questions were generally difficult, and the questions were presented in easy-to-understand language. The students' responses will serve as input for the researcher to revise the assessment for improvement. Here is the questionnaire sheet with the results of the small-scale test:

Table 4. The small-scale test results involving 12th-grade students

No	Question	Students' Answer			
		Strongly Agree	Agree	Disagree	Strongly Disagree
1	The questions are presented in easy-to-understand language.	2	4		
2	The instructions and information in the questions are clear and not confusing.	2	3	1	
3	The questions have a difficulty level appropriate for my abilities.		5		1
4	The questions give me the opportunity to think critically and creatively.	3	3		
5	The questions help me understand the concepts in the material taught.		6		
6	The questions are relevant to real-life situations or the application of learned concepts.	2	4		
7	The illustrations or images in the questions help me understand the questions asked.	4	1	1	
8	The questions are presented with a neat and attractive appearance.	4	2		
	Overall Total (%)	**26%**	**69%**	**3%**	**2%**

After the small-scale assessment trial was conducted, the final revision stage aimed to refine the instrument based on the findings from the trial results. This step involves an in-depth analysis of the data obtained, including identifying shortcomings in the test items, instructions, or presentation format of the assessment. Feedback from trial participants, including aspects of technical difficulty and the alignment of questions with assessment objectives, will be considered by the researcher in the revision. The final revision also considers the consistency between the evidence generated and the expected inferences, in accordance with the principles of *Evidence-Centered Design* (ECD). If there are any changes, they will be adjusted based on the validation results and clarity of the instrument, so that it can be implemented optimally in the next phase or for wider use.

CONCLUSION

This research successfully produced a dynamic fluid assessment instrument based on *Evidence-Centered Design* (ECD) with very high validity (Aiken's V = 0.834). The developed instrument is capable of comprehensively measuring learning achievement, particularly in representation skills, data analysis, evidence-based explanation, and information evaluation. This assessment development is expected to assist various parties as literacy material and an example in designing a good assessment, of course using the principles of *Evidence-Centered Design* (ECD). And the test results show that this assessment is able to measure students' abilities in accordance with learning outcomes. The implication is that this assessment can serve as a reference for teachers and educators in developing instruments that are not only valid and reliable, but also authentic, relevant, and supportive of improving the quality of physics learning in schools.

REFERENCES

Abidin, Z. (2007). Analisis kebutuhan pembelajaran dan analisis pembelajaran dalam desain sistem pembelajaran. *SUHUF*, (19) 1.

Affandy, H., Aminah, N. S., & Supriyanto, A. (2019). Analisis keterampilan berpikir kritis siswa pada materi fluida dinamis di SMA Batik 2 Surakarta. *Jurnal*

Materi dan Pembelajaran Fisika (JMPF). 9 (1). https://doi.org/10.20961/jmpf.v9i1.31608

Amaliah, N., Sari, T. M., Nurdin, G. M., Mustaqim., Salewangeng, A., & Akbar, M. N. (2023). Pengembangan instrumen asesmen sebagai upaya peningkatan keterampilan abad 21. *SAINTIFIK: Jurnal Matematika, Sains, dan Pembelajarannya*, *9*(1). https://doi.org/10.31605/saintifik.v9i1.390

Boy, N. D. (2022). Pengembangan asesmen pembelajaran bahasa dan sastra indonesia. *Jurnal Panca Widha, 1*(1), 18-35.

Branch, R. M. (2009). Instructional design: The ADDIE Approach. *Springer*. https://doi.org/10.1007/978-3-319-19650-3_2438

Damayanti, S. A., & Untari, E. (2024). Kesesuaian asesmen dengan tahapan perkembangan dan kemampuan kognitif peserta didik kelas 1 mata pelajaran bahasa indonesia. *Journal of Language Literature and Arts, 4*(9), 966-971. https://doi.org/10.17977/um064v4i92024p966-971

Hajati, K., & Amaliah, N. (2022). Teacher's delivery format and ADDIE model to adjust the online learning system during the pandemic period. *Indonesian Research Journal in Education (IRJE), 6*(1), 91-104. https://doi.org/10.22437/irje.v6i1.14649

Hanum, S. A. (2021). Analisis effect–size pengaruh bahan ajar fisika dan ipa terpadu terhadap hasil belajar siswa. *Jurnal Penelitian dan Pembelajaran Fisika, 7*(2), 144-153. 10.24036/jppf.v7i2.111741

Hendryadi. (2017). Validitas isi: Tahap awal pengembangan kuesioner. *Jurnal Riset Manajemen dan Bisnis (JRMB)*. *2*(2). 10.36226/jrmb.v2i2.47

Huitt, W., Hummel, J., & Kaeck, D. (2001). *Assessment*, measurement, evaluation, and research. *Educational Psychology Interactive*.

Kadir, A. (2015). Menyusun dan menganalisis tes hasil belajar. *Al-TA'DIB: Jurnal Kajian Ilmu Kependidikan*, *8*(2), 70-81.

Kasdriyanto, D. Y., & Wardana, L. A. (2021). Pengembangan media scrapbook berbasis picture and picture berorientasi wawasan kebangsaan. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, *6*(1), 271-278. 10.31004/obsesi.v6i1.1255

Kemdikbud. (2022). Capaian pembelajaran mata pelajaran fisika fase E – fase F untuk SMA/MA/Program paket C. *Badan Standar, Kurikulum, Dan Asesmen Pendidikan Kementerian Pendidikan, Kebudayaan, Riset, Dan Teknologi.*

Kurnia, T. D., Lati, C., Fauziah, H., & Trihanton, A. (2019). Model ADDIE untuk pengembangan bahan ajar berbasis kemampuan pemecahan masalah berbantuan 3d pageflip. *Prosiding Seminar Nasional Pendidikan Matematika (SNPM), 1(1)*, 516-525.

Maemonah. (2018). Asesmen pembelajaran. PGMI Press: Yogyakarta

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (1999). Evidence-centered assessment design. Educational Testing Service.

Monteiro, V., Mata, L., & Santos, N.N. (2021). Assessment conceptions and practices: perspectives of primary school teachers and students. *Frontiers in Education*. https://doi.org/10.3389/feduc.2021.631185

Munaroh, N. L. (2024). Asesmen dalam pendidikan: Memahami konsep, fungsi dan penerapannya. *Dewantara: Jurnal Pendidikan Sosial Humaniora*, *3*(3), 281-297. https://doi.org/10.30640/dewantara.v3i3.2915

Nabil, N. R. A., Wulandari, I., Yamtinah, S., Ariani, S. R. D., & Ulfa, M. (2022). Analisis indeks Aiken untuk mengetahui validitas isi instrumen asesmen kompetensi minimum berbasis konteks sains kimia. *Jurnal Penelitian Pendidikan*, *25*(2), 184-191. 10.20961/paedagogia.v25i2.64566

Nawawi, S., & Wijayanti, T. F. (2018). Pengembangan asesmen biologi berbasis keterampilan berpikir kritis terintegrasi nilai islam. *Jurnal Inovasi Pendidikan IPA*, *4*(2), 136-148. 10.21831/jipi.v4i2.21265

NGSS Lead States. (2013). Next generation science standards: For states, by states. Washington, DC: The National Academies Press.

Norlaila., Saadi, P., Winarti, A., & Leny. (2024). Development of e-module sets-based approach assisted with flip html5 media to improve independent learning and learning knowledge outcomes on chemical equilibrium material. *Quantum: Jurnal Inovasi Pendidikan Sains*, *15*(1), 122-139. http://dx.doi.org/10.20527/quantum.v15i1.17046

Puteri, A. N., Yoenanto, N. H., & Nawangsari, N. A. F. (2023). Efektivitas asesmen autentik dalam pembelajaran. *Jurnal Pendidikan Dan Kebudayaan*, *8*(1), 77-87. https://doi.org/10.24832/jpnk.v8i1.3535

Rahma, A. N. (2012). Pengembangan perangkat pembelajaran model inkuiri berpedekatan sets materi kelarutan dan hasil kali kelarutan untuk menumbuhkan keterampilan berpikir kritis dan empati siswa terhadap lingkungan. *Journal of Educational Research and Evaluation*, *1*(2), 133-138.

Sarwi., Isnaeni, W., & Ellianawati. (2019). Asesmen sains berorientasi next generation sains standards (NGSS). Unnes Press: Semarang.

Septyana, R., Nuzula, S. F., & Gusanti, Y. (2024). Peningkatan asesmen formatif melalui pemanfaatan media gamifikasi gimkit terhadap hasil belajar peserta didik SMPN 24 Malang. *Jurnal Integrasi dan Harmoni Inovatif Ilmu-Ilmu Sosial*, *4*(4), 7-7. https://doi.org/10.17977/um063v4i4p7

Sirait, J. (2024). Merancang asesmen fisika menggunakan evidence-centered design (ECD) dan three-dimensional learning protocol (3D-LAP). *Jurnal Pendidikan Informatika dan Sains,* *13*(2), 165-175. https://doi.org/10.31571/saintek.v13i2.8228

Soeyono, Y. (2014). Pengembangan bahan ajar matematika dengan pendekatan open-ended untuk meningkatkan kemampuan berpikir kritis dan kreatif siswa SMA. PYTHAGORAS: *Jurnal Matematika dan Pendidikan Matematika*, *9*(2), 205-218. https://doi.org/10.21831/pg.v9i2.9081

Sugiyono. (2016). Metode penelitian kuantitaf kualitatif dan R&D. Bandung: Alfabeta.

Voogt, J & Roblin, N.P. (2010). 21st century skills discussion paper. University of Twente.

Zural, M. M., Utami, D., Abizar., & Nurhadi. (2024). Development of a practicum guide for reproduction material in plants integrated with modified free inquiry for class XI SMA/MA. *Quantum: Jurnal Inovasi Pendidikan Sains*, *15*(2), 186-196. http://dx.doi.org/10.20527/quantum.v15i2.19050