ISSN 2987-4270
PLE
Physics Learning and Education
Vol. 3 No. 4 (2025)

DESIGN OF AN ELECTRONIC ASSESSMENT TO ASSESS STUDENTS' CRITICAL THINKING ABILITIES ON STATIC FLUID MATERIAL

Amelia Zelika Herawan[1], Fatni Mufit[1*], Festiyed[1],Selma Riyasni[1]

[1] *Department of Physics, Universitas Negeri Padang, Jl. Prof. Dr. Hamka Air Tawar Padang 25131, Indonesia*
Corresponding author. Email: fatni_mufit@fmipa.unp.ac.id

ABSTRACT

Assessment of students' critical thinking abilities is a crucial component of 21st-century learning. This study aims to design an electronic assessment instrument to evaluate students' critical thinking skills in the topic of static fluid using a summative assessment approach. The instrument consists of twenty multiple-choice questions developed based on Ennis's critical thinking indicators. The research employed the Plomp model, which includes three stages: preliminary research, prototyping, and assessment development. Validation by three experts indicated a high level of validity (Aiken's V = 0.83), while the practicality test showed positive responses from both students (83.9) and teachers (88.4). The reliability coefficient was 0.75 (high category), with items categorized as easy to moderate in difficulty and adequate in discrimination power (0.36). This study contributes by providing a digital-based assessment tool that supports the measurement of higher-order thinking in physics education. However, the study was limited by a small sample size and testing conducted only on static fluid material. Future research is recommended to expand the sample and apply the instrument across different physics topics to enhance its generalizability.

Keywords: *electronic assessment, critical thinking, static fluid.*

I. INTRODUCTION

The Industrial Revolution 4.0 era is marked by the rapid development of science and technology[1]. This development creates new challenges for the world of education[2]. 21st century learning demands mastery of core competencies such as collaboration, creativity, communication, and critical thinking[3]. Among these competencies, critical thinking skills are an essential skill that must be possessed to face the changes and complexities of the modern world[4]. Thus, students must be able to possess 21st-century critical thinking skills, be able to digest information efficiently and draw conclusions based on data, and be in accordance with the use of the independent learning curriculum in schools.

In using the independent curriculum, which emphasizes contextual learning that includes critical thinking skills, this curriculum gives students the opportunity to improve their critical thinking skills by studying concepts in real-world situations[5]. This approach aims to create a learning environment that not only pursues academic achievement but also enhances students' ability to understand, analyze, and evaluate ideas in a variety of contexts[6]. So the independent curriculum gives students the freedom to learn according to their interests and talents, while developing critical thinking skills.

According to Miele & Wigfield, critical thinking skills include the ability to connect ideas and knowledge to draw logical and reliable conclusions. Ennis & Weir explain that critical thinking is a logical and rational reasoning process in making belief-based decisions. This ability helps individuals analyze and solve problems intellectually. Therefore, students need to master critical thinking skills with the support of teachers' active role in their development[7]. Learning physics plays an important role in developing critical thinking skills that can be used to solve practical problems in everyday life.

Learning Physics is not a subject that requires memorization, but rather demands mastery of the material in a factual, conceptual, principled and procedural manner[8]. Understanding concepts and a positive attitude towards learning physics are prerequisites for success in learning physics[9]. Developing students' abilities in learning physics is one of the keys to success in increasing their ability to adapt to technology, improving problem-solving skills, and improving critical thinking skills[10]. Fluids are an important aspect of everyday life,

both in liquid and gas form. One branch of physics that studies fluids is fluid statics, namely fluids that are still or not experiencing flow [11]. This static fluid can help teachers assess students' critical thinking skills due to its complex nature and the thought processes involved. The concept of static fluid can train and equip students with an understanding of critical thinking[12].

Students' critical thinking skills are still relatively low, especially in fluids. According to Nasution's 2021 research, students' critical thinking skills were deemed very poor[13]. Students' critical thinking skills in Ardiyanti and Nuroso's 2021 research were in the low range[14]. According to Nurjannah's 2022 research, students' critical thinking skills in fluid material are in the low range[15]. Naila's 2023 study on low/medium level students' critical thinking skills with thermodynamics content [16]. Amelia and Chusni's 2024 research also stated that students' critical thinking skills in physics learning were still relatively low[13]. The five journal and article literature studies have similarities in the form of using critical thinking indicators according to Ennis 2011[14]. The advantage of this research over previous research is the use of electronic assessment with critical thinking indicators.

Based on observations, the assessments used by teachers in schools do not yet include critical thinking indicators, so adjustments are needed to be able to measure students' critical thinking abilities effectively. To address these needs, researchers developed an electronic assessment specifically designed to assess critical thinking skills, especially on the concept of static fluids[15]. Electronic assessments offer efficiency and flexibility because they can be accessed online, saving correction time, and allowing for more accurate data analysis[16]. This innovation helps teachers monitor student learning progress objectively, is environmentally friendly, efficient in storing and managing results, and is easy to integrate with educational technology[17].

Based on the background that has been explained, the author is interested in developing an electronic assessment to assess students' critical thinking skills on static fluid material. This electronic assessment is a technology-based digital assessment that can be accessed via electronic devices. Therefore, the research entitled "Electronic Assessment Design to Assess Students' Critical Thinking Skills on Static Fluid Material" was published in the journal Science.

II. METHOD

This study uses a Research and Development (R&D) approach (Sugiyono, 2019) [18] with the Plomp development model to design and publish electronic-based summative assessments containing critical thinking indicators on statistical fluid material. The developed product is implemented through the Wizer.me platform, with the aim of improving students' critical thinking skills in physics learning. The Plomp (2013) development model consists of three main stages, namely the preliminary research stage (preliminary research), the prototyping stage (prototyping phase), and the assessment stage (assessment phase) [19].

The preliminary research phase aimed to identify the needs and theoretical basis for developing assessment instruments. A needs analysis was conducted through interviews with three high school physics teachers to identify obstacles in the learning and assessment process for static fluids, particularly those related to the application of critical thinking indicators. Furthermore, a literature review of various scientific journals was conducted to determine the conceptual framework, critical thinking indicators, and design principles for developing electronic-based assessments. The results of this phase included design specifications and content indicators, which served as the basis for developing the assessment prototype.

The prototyping stage encompasses the design, validation, and refinement processes of the product. The initial design included the development of test items based on critical thinking indicators according to Ennis (2011), tailored to the objectives of static fluid learning. Furthermore, the design included an interface, instructions, and digital features using the Wizer.me platform. Formative evaluation was conducted through several steps: self-evaluation, expert review, and revision. In the self-evaluation stage, the researcher reviewed the initial prototype to ensure completeness of content, technical accuracy, and clarity of language. Next, in the expert review stage, three physics education lecturers assessed the content, construction, and language using a product validation sheet. The validity of the test items was analyzed using Aiken's V formula, while inter-rater agreement was examined through assessment consistency to ensure the reliability of the validation results. Based on expert input, revisions were made until a validated assessment prototype was obtained and ready for limited testing.

The assessment phase aims to evaluate the practicality and empirical quality of the validated product through a limited trial. The trial was conducted on nine students and three physics teachers from a high school in Padang. The limited number of participants was adjusted to the character of the initial development research stage, so the results were not intended to be generalized, but rather to assess the feasibility of the product. The instruments used included observation sheets, student response questionnaires, and test results that were analyzed to assess the practicality and quality of the test items. Data were analyzed descriptively, including validity using Aiken's V, reliability using Cronbach's Alpha with the criteria of rxy count > rxy table declared

reliable, as well as calculation of the level of difficulty and discriminatory power to assess the quality of each test item. Because this study focused on product development and validation, inferential statistical analysis was not applied, and the results of the analysis focused on the feasibility, validity, and reliability of the developed assessment.

Based on the results of n experts' evaluation of an item, the content validity coefficient is determined using the V-Aiken formula (Aiken, 1985). The following is the V-Aiken formula used:

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

Katerangan :
$l_0 = lowest\ validity\ assessment\ score$
$c = highest\ validity\ assessment\ number$
$r = the\ number\ given\ by\ the\ validator$

The results of the Aiken's V formula can be interpreted as falling between 0 and 1. Table 1 below displays the Aiken's V validity index.

.

Table 1. Validity Assessment Index

Intervals	validity category
$\leq 0{,}4$	not enough
$0{,}4 < V \leq 0{,}8$	Currently
$0{,}8 < V$	Height

III. RESULTS AND DISCUSSION

A. Research Results

1. Needs Analysis

In the initial stage of the research, a preliminary study was conducted, analyzing the use of assessment in educational institutions. The preliminary research on educators was conducted through interviews with three physics teachers at SMAN 8 Padang. In the preliminary research stage, a preliminary study was conducted in the form of an analysis of assessment practices in schools. This preliminary study involved interviews with three physics teachers at SMAN 8 Padang. The purpose of these interviews was to understand how assessment is implemented in schools.

From interviews with three physics teachers, the following results were obtained: (1) assessment questions in schools still do not have critical thinking indicators, (2) summative assessment questions are still in the form of mathematics and calculations, (3) the use of assessments in schools is less interesting and innovative, (4) students' critical thinking skills are still low, (5) the use of electronic assessments is not optimal, especially in static fluid material. Based on the initial analysis of teachers, it was concluded that the creation of electronic assessments to measure students' critical thinking skills in static fluid material is quite important.

After that, a literature study was conducted on articles/journals regarding the level of students' critical thinking skills. The results of this literature study found that high school students' critical thinking skills are still relatively low. For example, a study by Nasution et al. (2021) stated that students' critical thinking levels are in the category with the lowest proportion, 39.49% [18]. According to a research by Ardiyanti & Nuroso (2021), 55.6% of pupils had poor critical thinking abilities[19]. According to a research by Naila (2024), 46.76% of students expressed having poor to moderate critical thinking abilities[20]. According to a research by Nurjanah et al. (2022), pupils' critical thinking abilities fall into the low range[21], and a study by Amelia & Chusni (2024) found that students' critical thinking skills are in the low category with a percentage of 37.21% [22]. In accordance to the findings of the examination of several publications and papers, students' capacity for critical thinking in the study of physics remain comparatively poor.

According to the requirements analysis, a specially created test is requested to gauge and enhance students' critical thinking abilities. A solution must be created to address the issues raised by this preliminary investigation. This involves establishing an evaluation instrument to help instructors gauge their the capacity for critical thinking in learners in relation to the subject of static fluids. This assessment is designed in essay format using the Google Sites and Wizer.me platforms.

1. Development Stage *(Development Or Prototyping Phase)*
   a. Prototype Design

Assessment product design, self-evaluation, and expert validation are the outcomes of this design phase. To evaluate students' critical thinking abilities on static fluid material, an electronic evaluation was developed as the product design. The intended assessment is a summative one, which is conducted at the conclusion of a learning session to gauge how well students have met learning goals and/or Learning Outcomes (CP). At the conclusion of a learning session, the summative assessment findings serve as the foundation for assessing how well students have met their learning objectives or learning outcomes (CP). This assessment is made in the form of an essay developed based on critical thinking indicators according to Ennis (2011) [22]. This assessment is designed using the Google.sites and Wizer.me websites as shown in Figures 1(a) 1(b) and 2(a) 2(b).

Figure 1(a). Home page

Figure 1(b). Content section

Figures 1(a) and 1(b) represent assessment designs created using the Google Sites website. The use of Google Sites in this assessment allows students and teachers easy access to the assessment. Teachers can create and manage the website quickly and easily without requiring coding skills. Students can also easily access the assessment for free and connect directly to wizer.me and other Google Docs products such as gDrive and Canva. The use of Google Sites also aims to make the assessment more attractive and presentable, as it includes user instructions and an assessment outline from Canva, which has an attractive appearance.

Figure 2(a).Discourse display on wizer.me

Figure 2(b).Answer sheet display on wizer.me

Figures 2(a) and 2(b) represent the summative assessment design using the wizer.me website. Wizer.me was chosen because it is web-based and does not require downloading; it can be accessed by simply opening Google, Chrome, or Firefox via https://wizer.me.com. Educators and students can use it flexibly. Wizer.me has an attractive interface, allowing for the addition of audio, images, and videos, as well as direct student response columns [23]. The advantages of wizer.me's e-assessments include the ability to be designed as attractively as possible with a variety of themes provided on the website; the wide variety of features tailored to individual needs; the assessments can be accompanied by images, audio, or video to support student understanding; students can access them via smartphone, tablet, or laptop; and the entire assessment process, from completing the assessment to collecting answers, can be done online. Therefore, the electronic assessments are designed using wizer.me as the medium and Google Sites as additional media.

b. Formative Evaluation

a) Self Evaluation Results

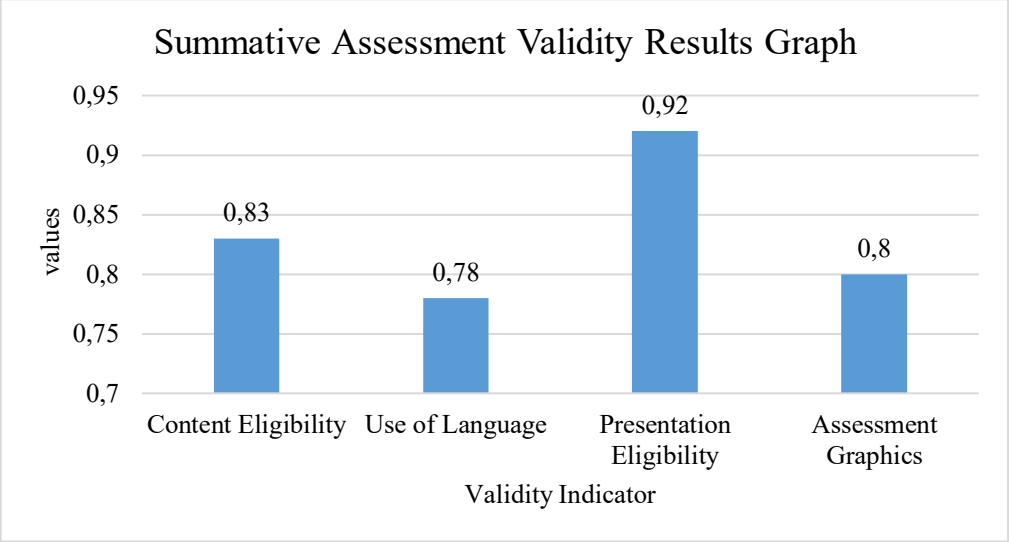
After the assessment prototype was developed, a formative evaluation was conducted in two stages: self-evaluation and expert review. Self-evaluation is an assessment conducted by researchers to assess the completeness and suitability of the electronic assessment developed to measure students' critical thinking skills on the topic of static fluids. This self-evaluation stage was conducted before the product was reviewed by a validator for validation. At this stage, researchers read the material, checked for completeness, revised errors, and added sections that felt incomplete. Indicators assessed in the self-evaluation include three aspects: assessment structure, language, and assessment graphics.

b) Results Of Content Validity Testing By Experts (Expert Review)

Experts carried out the expert review or validity test phase. Three UNP lecturers validated the electronic test used to gauge students' critical thinking abilities on static fluid material. The validation's findings are used as a reference to update the product and analyze the viability of the created evaluation. Four components make up the product validation tool: assessment visuals, language use, presentation viability, and material content. Data analysis was done to ascertain the validity of the electronic evaluation after the validator supplied a score scale. The four elements' validation scores were examined based on the assessment tool that was employed. For each validation aspect indication, Aiken's formula was used to examine validity test results based on assessment aspects. There are 18 evaluation markers across four elements of the assessment.

Three UNP physics instructors validated an electronic test to gauge students' critical thinking abilities on the subject of static fluids. The validation's findings served as a roadmap for the product's revision and viability assessment. Four components made up the product validation tool: assessment visuals, language use, presentation feasibility, and content feasibility. The three experts' validation results, including feedback and suggestions for revision, are presented in Figure 1.

Graph 1. Results of the validity of the summative assessment



Based on the results of the assessment graph, of the 18 indicators analyzed, the four validity aspects had an overall average value of 0.83, which is included in the high category. For the content appropriateness indicator, the data analysis results obtained an average of 0.83, which is included in the high category. For the language use indicator, the data analysis results were 0.78, which is included in the medium category. Meanwhile, the presentation appropriateness indicator obtained a data analysis of 0.92, which is categorized as high. Then, for the assessment graphic indicator, the data analysis obtained a data analysis of 0.8, which is included in the high category. Of the four aspects, the lowest average value is for the language use indicator and the indicator with the highest average value is located in the presentation appropriateness indicator. From the analysis of all aspects of this validation indicator, it shows that in general the designed electronic assessment is valid.

Based on the assessment validation results, the validator also provided several suggestions and input. The validator provided suggestions and input so researchers could improve the assessment. These suggestions and input will be considered in the revised assessment. The validator's suggestions and comments can be seen in Table 2.

Table 2. Validator suggestions

Validator	Suggestion
Validator 1	1. The discourse presented must be contextual. 2. Ensure that the images provided are in the surrounding environment.
Validator 2	1. Align the questions with the question indicators. 2. Align the question language with the learning objective indicators.
Validator 3	1. Use the correct and proper KKO. 2. Correct the statement. 3. Use the source in each image.

Advice provided by expert validators or lecturers during the assessment validation process plays a crucial role in ensuring the quality of the developed assessment instruments. The primary objective of this validation is to ensure that each item or assessment task aligns with the desired learning objectives. Through careful assessment, validators evaluate various aspects of the instrument, such as the appropriateness of the material, clarity of language, level of difficulty, and coverage of cognitive domains. The advice provided not only helps correct technical deficiencies but also improves the validity and reliability of the instrument, ensuring that the assessment accurately measures student competencies. Furthermore, the involvement of expert validators is crucial to avoid bias or inaccuracy that could impact assessment results. Therefore, this validation process is an integral part of developing professional, accountable, and high-quality assessments.

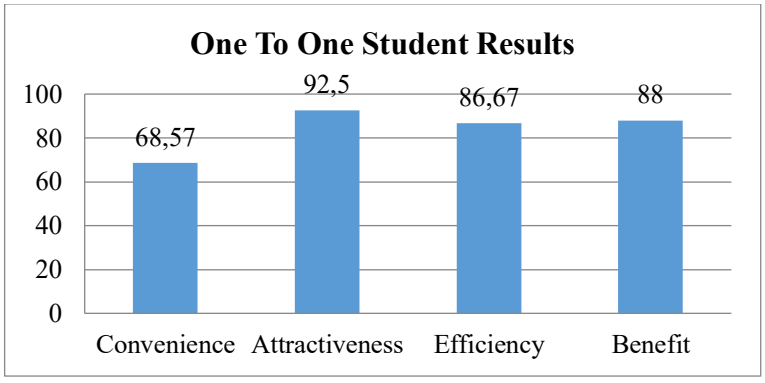
c) One To One Evaluation

The one-to-one test is a practicality test conducted with three students and a teacher. Through the one-to-one trial, researchers were able to determine the practicality of the electronic assessment for assessing students' critical thinking skills in static fluid motion. The goal was to assess its practicality for both students and teachers. Therefore, researchers need to improve the electronic assessment for assessing students' critical thinking skills in static fluid motion.

A one-to-one test was conducted on three students of SMAN 8 Padang to assess the practicality of the developed electronic assessment instrument, specifically in assessing critical thinking skills on static fluid material. The assessment was conducted using a practicality questionnaire that included four indicators, namely ease, attractiveness, efficiency, and usefulness. The ease indicator consisting of seven statement items obtained an average score of 68.57 and was included in the practical category. The attractiveness indicator consisting of eight statement items produced an average score of 92.50 with a very practical category. The usefulness indicator, which had five statement items, had an average score of 88.00, while the efficiency indicator, which had four statement items, received an average score of 86.67. Both indicators fell into the extremely practical group. With an overall score of 83.9 across the four parameters, this electronic evaluation tool fell into the extremely practical category. Therefore, based on the results of the one-on-one exam given to students, it can be said that the electronic assessment to gauge critical thinking abilities on static fluid material was deemed very practical. Table 2 and Graph 2 present the findings.

Table 2. One to one student results

Statement Item	Convenience	Attractiveness	Efficiency	Benefit
1	66,67	93,33	80,00	86,67
2	66,67	93,33	86,67	86,67
3	66,67	93,33	86,67	86,67
4	73,33	93,33	93,33	86,67
5	73,33	93,33	-	93,33
6	66,67	93,33	-	-
7	66,7	93,33	-	-
8	-	86,67	-	-
Rata-rata	**68,57**	**92,50**	**86,67**	**88,00**



Graph 2. One to one student results

Then, a one-to-one test was conducted on teachers using a practicality instrument questionnaire for 3 physics classes at SMAN 8 Padang to test the practicality of the designed electronic assessment. The assessment was conducted using a practicality questionnaire that included four indicators, namely ease, attractiveness, efficiency, and usefulness. With an average score of 90.48, the seven statement items that made up the easiness indicator fell into the extremely practical category. Eight statement items made up the attractiveness indicator, which yielded an average score of 85.00 in the extremely practical category. The usefulness indicator, which had five statement items, had an average score of 93.33, while the efficiency indicator, which had four statement items, received an average score of 85.00. Both indicators fell into the extremely practical category. With an overall score of 88.4 across the four parameters, this electronic evaluation tool fell into the extremely practical category. Table 3 and Graph 3 present the findings.

Table 3. Results of one to one teacher

Statement Item	Convenience	Attractiveness	Efficiency	Benefit
1	93,33	80,00	86,67	100,00
2	86,67	80,00	93,33	100,00
3	93,33	86,67	86,67	80,00
4	93,33	93,33	73,33	86,67
5	80,00	73,33	-	100,00
6	86,67	73,33	-	-
7	100,00	93,33	-	-
8	-	100,00	-	-
Rata-rata	**90,48**	**85,00**	**85,00**	**93,33**

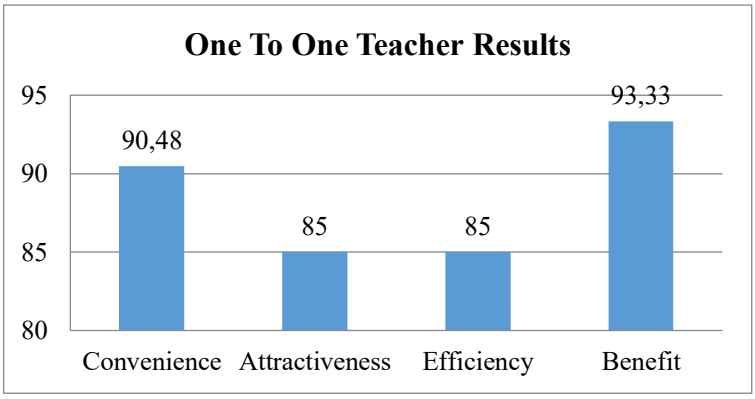


Chart 3. One to one teacher results

From table 3 and graph 3 it is shown that the instrument developed has a high level of practicality and is in accordance with the criteria set based on tests on teachers. Therefore, based on the results of the one-on-one exam given to instructors, it can be said that the electronic assessment to gauge critical thinking abilities in static fluid material was deemed very practical

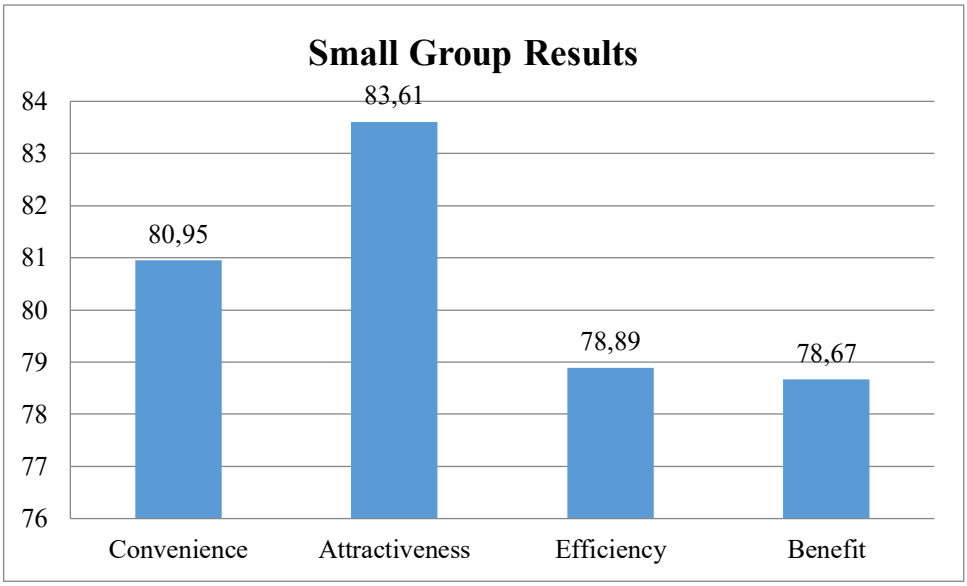
d) Small group

After a one-to-one trial, Nine students representing low, medium, and high ability levels participated in a small group trial of the electronic exam used to evaluate students' critical thinking abilities in static fluids. The trial in small groups was conducted using a practicality questionnaire with nine students from SMAN 8 Padang. Through the small group practicality sheet, The usefulness of the electronic test for evaluating students' critical thinking abilities in static fluids was ascertained by the researchers. Four primary criteria were used to distribute the questionnaire: advantages, convenience, attractiveness, and efficiency. Students were required to complete 24 questions based on these four indications.

The findings from the analysis of the practicality questionnaire in the small group trial indicate that the developed electronic assessment instrument falls into the practical category. The assessment was conducted based on four indicators: ease, attractiveness, efficiency, and usefulness. The ease indicator, consisting of seven statements, obtained an average score of 80.95, which is included in the practical category. The attractiveness indicator obtained an average score of 83.61 and is included in the very practical category. Furthermore, the efficiency indicator produced an average of 78.89, and the usefulness indicator obtained an average of 78.67; both are included in the practical category. The instrument is classified as practical as the average of the four indications was 80.53 overall. The electronic test to gauge critical thinking abilities in static fluid content is therefore deemed feasible and appropriate for use in the following phase based on the outcomes of the small group experiment. Table 4 and Graph 4 present the findings.

Table 4. Small Group Results

Statement Item	Convenience	Attractiveness	Efficiency	Benefit
1	75,56	84,44	73,33	75,56
2	80,00	84,44	77,78	82,22
3	86,67	80,00	80,00	86,67
4	80,00	84,44	84,44	75,56
5	80,00	84,44	-	73,33
6	80,00	77,78	-	-
7	84,44	84,44	-	-
8	-	88,89	-	-
Rata-rata	**80,95**	**83,61**	**78,89**	**78,67**



Graph 4. Small group results graph

From table 4 and graph 4 it is shown that the instrument developed is practical and in accordance with the criteria set out in a limited trial using 9 students. This research is still on a small scale and focuses on presenting test items.

c. Assessment Phase (Assessment Phase)
   a. Trial Questions (Fielt Test)
      a) Construct Validity

Table 5. Results of construct validity analysis

Question no	R_{XY}	Description	Interpretation
1.a	0,480	Valid	Currently
1.b	0,679	Valid	Height
1.c	0,736	Valid	Height
1.d	0,754	Valid	Height
1.e	0,877	Valid	Very High
2.a	0,840	Valid	Very High
2.b	0,465	Valid	Currently
2.c	0,787	Valid	Height
2.d	0,783	Valid	Very High
2.e	0,785	Valid	Height
3.a	0,800	Valid	Height
3.b	0,658	Valid	Height
3.c	0,702	Valid	Height

3.d	0,500	Valid	Currently
3.e	0,688	Valid	Height
4.a	0,831	Valid	Very High
5.a	0,735	Valid	Height
5.b	0,903	Valid	Very High
6.a	0,875	Valid	Very High
6.b	0,737	Valid	Height

From Table 5, the results of the construct validity analysis show that all assessment questions on the critical thinking indicator for static fluid material have been declared valid. The average value of the validity coefficient obtained is 0.73, which according to the interpretation of the validity criteria, falls into the high group.

b) Reliability Results

Table 6. Reliability Analysis Results

R_{11}	**Description**
0,95	Reliabel

From Table 6, the results of the reliability analysis show that all assessment questions on the critical thinking indicator for the static fluid material have been declared reliable. The R11 value obtained is 0.95, where the calculated RXY > Rxy table. Therefore, it can be stated that the critical thinking assessment questions for the static fluid material have been declared reliable and are suitable for assessing students' critical thinking abilities.

c) Difficulty Level Results

Table 7. Results of the analysis of the level of difficulty

Question no	P	Criteria
1.a	0,86	Easy
1.b	0,66	Currently
1.c	0,69	Currently
1.d	0,69	Currently
1.e	0,68	Currently
2.a	0,70	Currently
2.b	0,69	Currently
2.c	0,67	Currently
2.d	0,67	Currently
2.e	0,67	Currently
3.a	0,66	Currently
3.b	0,69	Currently
3.c	0,68	Currently
3.d	0,69	Currently
3.e	0,68	Currently
4.a	0,69	Currently
5.a	0,47	Currently
5.b	0,70	Currently
6.a	0,65	Currently
6.b	0,67	Currently

According to Table 7. The examination of 20 critical thinking indicator assessment questions on static fluid material revealed that, on average, the questions fall into the moderate difficulty range, indicating that students can typically complete them successfully.

d) Differential Power Results

Table 8. Results of the analysis of discriminant power

Question no	D	Criteria
1.a	0,23	Enough
1.b	0,26	Enough
1.c	0,37	Enough
1.d	0,31	Enough
1.e	0,40	Enough
2.a	0,40	Enough
2.b	0,43	Good
2.c	0,37	Enough
2.d	0,49	Good
2.e	0,49	Good
3.a	0,40	Enough
3.b	0,31	Enough
3.c	0,29	Enough
3.d	0,31	Enough
3.e	0,26	Enough
4.a	0,43	Good
5.a	0,29	Enough
5.b	0,49	Good
6.a	0,40	Enough
6.b	0,34	Enough

Based on Table 8, the results of the discriminant power data analysis show that of the 20 questions given to students, the average discriminant power analysis was at 0.36, which is included in the sufficient category. This indicates that the summative assessment instrument is ready to be used, but the quality of items in the sufficient category can still be improved to have more optimal discriminating power.

Overall, the developed media demonstrated high levels of validity, practicality, and effectiveness. However, effectiveness testing was limited to a small scale and did not involve diverse student backgrounds. Further research with a broader sample size is needed to ensure consistency of results across diverse learning contexts.

B. Discussion

The Industrial Revolution 4.0, which emphasizes the importance of technology in education, is the basis for the creation of this computerized evaluation. Critical thinking skills are seen as a basic skill that must be possessed in the 21st century, just like reading and writing [24]. Humans in this century must also be able to think more critically and be able to keep up with rapid technological developments [25]. Therefore, a summative electronic test was created to highlight students' critical thinking skills.

Considering the literature study, it was found that students often have a low level of critical thinking in physics. To address this problem, an engaging and effective electronic assessment is needed that can directly measure students' critical thinking skills[26]. According to the findings of a preliminary study involving interviews with physics instructors at SMAN 8 Padang, teachers still rarely give exams to evaluate students' critical thinking skills for various reasons, including lack of time, lack of or minimal teacher training in making assessments, lack of innovation and references and examples relevant to static fluid material. So the researchers developed this electronic assessment to assist teachers in assessing students' critical thinking skills in static fluid material.

Validity and practicality findings indicate that e-assessment can help students understand the concept of static fluids more interactively. This is in line with constructivism theory which emphasizes the importance of active student involvement in the learning process. High assessment validity is supported by the suitability of the question content with critical thinking indicators according to Ennis and its relationship to real-life phenomena in everyday life. The authentic context of this assessment not only increases student engagement but also strengthens the relevance of the learning material. This is in line with research by [27] who stated that the contextual learning approach is effective in developing critical thinking skills because it provides a meaningful learning experience. Furthermore, the use of a systematic visual approach in presenting questions also supports

display validity, as graphic aspects have been shown to influence student focus and ability in electronic-based assessments.

For teachers, this e-assessment can be a tool to identify student understanding in real time. For students, this system helps them reflect on conceptual errors and correct them independently. It is important for critical thinking tests to use easily understood language so that the content of the questions is not hidden, because the quality of the questions depends heavily on how well the wording matches the indicators of critical thinking skills. Because the results of the electronic validity test for the critical thinking indicator of static fluids were of very good quality in terms of content, construction, and language, this evaluation has met the requirements of relevance and consistency [28]. Despite the positive results, implementing e-assessment in schools with limited ICT infrastructure is a challenge. Teachers need training support and facilities so that the use of e-assessment can be effective in all school contexts.

In terms of practicality, this electronic assessment demonstrates advantages in terms of ease of use, efficiency, appeal, and usefulness. Platforms such as Google Sites and Wizer.me allow for efficient and flexible question distribution, making it easier for teachers and students to carry out the evaluation process, even in a distance learning context. This finding aligns with research by Syafatullhalamah (2025), which also showed that e-assessment increases student engagement. However, unlike that research, this study focused on the concept of static fluids, which requires interactive visualization [29]. The assessment's appeal was also enhanced by the integration of interactive and visual elements. This aligns with the principle of visual literacy; appeal is a crucial component in developing meaningful e-learning. [30]. Furthermore, the presence of this assessment can encourage a more efficient and reliable evaluation process, thus emphasizing the importance of electronic assessments that support the development of critical thinking skills [22]. This study is limited by its small sample size and limited school context. Furthermore, the material tested only covered the topic of static fluids, so the results cannot be generalized to other physics materials. Thus, it can be concluded that the electronic assessment developed has met the criteria of validity and practicality, and has advantages in supporting the measurement of students' critical thinking abilities effectively, efficiently, and relevant to the needs of 21st century learning.

IV. CONCLUSION

The development of an electronic assessment design to measure students' critical thinking skills on static fluid material has gone through systematic validation and trial stages, resulting in a product that is valid, practical, and reliable. The high validity score (Aiken's V = 0.83), strong practicality evaluations from teachers and students, and high reliability indicate that the developed assessment instrument is feasible for classroom implementation. Beyond these statistical results, this study provides meaningful implications for educational practice. The electronic assessment serves as an effective tool to evaluate students' higher-order thinking skills, particularly critical thinking, in science learning. Its digital format supports efficiency, interactivity, and accessibility—important elements in modern education aligned with the demands of the Industrial Revolution 4.0 era. For policymakers and curriculum developers, the findings suggest the need to integrate electronic-based assessments into science curricula and teacher training programs to strengthen critical thinking evaluation. Future researchers are encouraged to expand this model to other physics topics or educational levels to enhance its generalizability and impact. The novelty of this study lies in its integration of critical thinking measurement within a technology-based assessment specifically designed for static fluid material, an area that has received limited attention in previous studies. This innovation contributes to the development of modern, data-driven assessment practices that foster 21st-century learning skills.

REFERENCES

[1] Sumartono and N. Huda, "Educational Management in Indonesia as an Implementation of the Triple Helix to Prepare for the Industrial Revolution 4.0," *J. Andi Djemma*, vol. 3, no. 1, pp. 74–79, 2020.

[2] S. Al Asadullah and Nurhalin, "The role of character education in shaping the critical thinking skills of Indonesia's young generation," *Kaisa J. Educators and Learning*, vol. 1, no. 1, pp. 12–24, 2021, [Online].

[3] M. R. A.P, "Critical Thinking Skills and Learning Independence Through the Flipped Classroom Learning Model with a STEM Approach," *Semin. Nas. Pascasarj. 2020*, no. 20, pp. 292–297, 2020, [Online].

[4] I. F. Yuzan and I. S. Jahro, "Development of e-LKPD Based on Guided Inquiry on the Topic of Chemical Bonding to Measure Students' Critical Thinking Skills," *J. Encyclopedia of Educators and Learning Innovation Saburai*, vol. 2, no. 01, pp. 54–65, 2022, doi: 10.24967/esp.v2i01.1598.

[5] W. Purwandari, I. Nur Safitri, and M. Mutiara Karimah, "Exploring the Nature of Mathematics Learning

in Elementary Madrasahs in the Context of the Independent Curriculum," *Indones. Res. J. Educ.*, vol. 4, no. 4, pp. 1045–1060, 2024.

[6] N. Kollo and O. A. Suciptaningsih, "Students' Critical Thinking Skills through the Implementation of the Independent Curriculum," *JIIP - J. Ilm. Educational Science.*, vol. 7, no. 2, pp. 1452–1456, 2024, doi: 10.54371/jiip.v7i2.3845.

[7] Z. Zulfawati, T. Mayasari, and J. Handhika, "The Effectiveness of the Problem-Based Learning Model Integrated STEM Approach in Improving the Critical Thinking Skills," *J. Penelit. Fis. dan Apl.*, vol. 12, no. 1, pp. 76–91, 2022, doi: 10.26740/jpfa.v12n1.p76-91.

[8] F. Mufit and S. Syamsidar, "Development of Four-Tier Multiple Choice Test Instrument to Identify Students' Concept Understanding of Newton's Law Material," *JIPF (Jurnal Ilmu Pendidik. Fis.*, vol. 7, no. 2, p. 134, 2022, doi: 10.26737/jipf.v7i2.2369.

[9] F. Mufit, "A Five-Level Multiple Choice Instrument to Assess Students' Understanding of Rectilinear Motion," pp. 250–265, 2023.

[10] F. D. Amanda, U. P. Dewi, F. Mufit, and F. Festiyed, "Influence of Essay Assessment on Student Competency Achievement in Science Learning: Literature Review," *J. Researcher. Educator. Science,*vol. 9, no. 9, pp. 539–549, 2023, doi: 10.29303/jppipa.v9i9.4994.

[11] B. Aji Kurniawan, S. Kusairi, P. Parno, and A. Suyudi, "Analysis of concept mastery in the static fluid sub-material, class XI students of SMAN 1 Lawang in 2017/2018," *J. Science and Learning,*vol. 1, no. 7, pp. 578–586, 2021, doi: 10.17977/um067v1i7p578-586.

[12] I. Fajria, D. H. Putri, and I. Setiawan, "Development of Student Worksheets (Lkpd) for Straight Motion Kinematics in High School Using Learning Cycle 7E to Train Students' Critical Thinking Skills," *Amplitudo J. Ilmu dan Pembelajaran Fis.*, vol. 2, no. 2, pp. 151–160, 2023, doi: 10.33369/ajipf.2.2.151-160.

[13] N. Amelia and M. M. Chusni, "Analysis of Critical Thinking Skills in Physics Learning on Renewable Energy Material," *BIOCHEPHY J. Sci. Educ.*, vol. 4, no. 1, pp. 248–252, 2024, doi: 10.52562/biochephy.v4i1.1114.

[14] R. Ennis, "Critical Thinking: Reflection and Perspective Part II," *Inq. Crit. Think. Across Discip.*, vol. 26, no. 2, pp. 5–19, 2011.

[15] S. Guswina and F. Mufit, "Four-Tier Multiple Choice Test Design on Harmonic Vibration Material to Identify Conceptual Understanding of Grade X SMA/MA Students," *J. Penelit. Pembelajaran Fis.*, vol. 6, no. 2, pp. 183–192, 2020, doi: 10.24036/jppf.v6i2.110873.

[16] M. Hasan, Research-Based Learning: Theoretical Basis, Planning, Implementation, and Evaluation.2022. [Online]. Available: https://tahtamedia.co.id/index.php/issj/article/view/166/166

[17] M. K. Ashari, S. Athoillah, and M. Faizin, "APPLICATION-BASED E-ASSESSMENT MODEL IN HIGH SCHOOL IN THE DIGITAL ERA: SYSTEMATIC LITERATURE REVIEW," vol. 6, no. 2, pp. 132–150, 2023.

[18] Sugiyono, Quantitative, Qualitative and R & D Research Methodology.2020.

[19] N. Plomp, Nieveen and E. Folmer, "Educational Design Research Educational Design Research," *Netherlands Inst. Curric. Dev. SLO*, pp. 1–206, 2013, [Online]. Available: http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=EJ815766

[20] N. Muna and S. Linuwih, "Analysis of Students' Critical Thinking Skills Using a Two-Tier Multiple Choice Test Based on HOTS on Thermodynamics Material," *Pros. Semin. Nas. Fis.*, vol. 35, pp. 307–318, 2024, [Online].

[21] S. Nurjanah, T. Djudin, and Hamdani, "Analysis of Students' Critical Thinking Skills on the Topic of Dynamic Fluids," *J. Educ. Dev.*, vol. 10, no. 3, pp. 111–116, 2022.

[22] N. Yuliani and F. Mufit, "Design and Validity of Electronic Assessments to Assess Students Critical Thinking Abilities on Circular Motion Materials," vol. 3, no. 1, pp. 80–93, 2025.

[23] R. NADAPDAP, M. RETHA, and E. Ernalida, "Developing Minimum Competency Assessment Questions for Level 4 Reading Literacy Using the Wizer.Me Application at SMP Negeri 41…," vol. 1, no. 1, pp. 194–210, 2022, [Online].

[24] J. D. Putriani and H. Hudaidah, "Implementation of Indonesian Education in the Era of the Industrial Revolution 4.0," *Edukatif J. Ilmu Pendidik.*, vol. 3, no. 3, pp. 830–838, 2021, doi: 10.31004/edukatif.v3i3.407.

[25] F. A. Asna, "Efforts to Improve Students' Critical Thinking Skills in Science Learning with a STEM-Pjbl Based Approach," *Proceeding Semin. Nas. IPA*, pp. 312–320, 2024, [Online].

[26] T. Nurazizah, Ramadhani, and Wasis, "Assessment of High School Students' Critical Thinking Skills on Static Fluids Through Assessment for Learning," *J. Rev. Educators and Teaching*, vol. 7, no. 3, pp. 11455–11466, 2024.

[27] F. E. Jasper-Abowei and E. C. Victor-Ishikaku, "Contextual learning approach: A tool for enhancing

critical thinking skills amongst learners',” *Cent. Asian J. Soc. Sci. Hist.*, vol. 4, no. 5, pp. 72–83, 2023.

[28] F. Mufit, “A Five-Level Multiple Choice Instrument to Assess Students' Understanding of Rectilinear Motion,” vol. 6, pp. 250–265, 2023.

[29] A. Syafaatussalamah and D. E. Salsabilla, “The Effectiveness of Using Digital Media in Increasing Elementary School Students' Learning Motivation,” *Al-Tarbiyah J. Islamic Education Science.*, vol. 3, no. 3, pp. 11–24, 2025, [Online].

[30] S. Nurfadhillah, D. F. Rizkiya, K. Waro, R. Handayani, and P. Ayu, “Application of Visual Learning Media in Mathematics Learning at Bina Bangsa Elementary School,” *Ed. J. Education and Sciences*, vol. 3, no. 2, pp. 253–263, 2021, [Online].