

Item analysis of a creative thinking skills test on dynamic fluids material

Brilliana Ghorbiy*, Parno, Ahmad Taufiq

Department of Physics, Universitas Negeri Malang, Cakrawala Street No.5 Malang, 65145, Indonesia

*Corresponding author, email: brilliana.ghorbiy.2303218@students.um.ac.id

Article History

Received: 16 July 2025

Revised: 18 April 2026

Accepted: 20 May 2026

Keywords

Creative thinking

Dynamic fluids

Item analysis

Test instrument

Abstract

Creative thinking skills are essential in physics learning because students are required not only to understand formulas, but also to generate ideas, solve contextual problems, and explain scientific phenomena creatively. However, most physics assessments still emphasize procedural calculations and have not comprehensively measured students' creative thinking skills, particularly in dynamic fluid learning. Therefore, this study aims to examine the quality of essay test items developed to assess students' creative thinking skills in the context of dynamic fluids. A descriptive quantitative approach was employed involving 120 students who had previously studied dynamic fluids. The developed instrument consisted of four contextual essay questions representing the indicators of fluency, flexibility, originality, and elaboration. Data analysis was conducted using SPSS version 30.0 through empirical validity testing, reliability analysis, item difficulty level, and discrimination index analysis. The results showed that all items had very high validity coefficients ($r = 0.827-0.946$; $p < 0.001$), indicating strong construct validity. The instrument also demonstrated very high reliability with a Cronbach's Alpha coefficient of 0.908. Furthermore, all items were categorized as having moderate difficulty levels (0.476-0.526) and good to very good discrimination indices (0.694-0.899). These findings indicate that the developed instrument is effective for measuring students' creative thinking skills in dynamic fluid learning. The novelty of this study lies in the development of a contextual essay-based assessment instrument specifically designed for dynamic fluid topics, where each item explicitly represents one indicator of creative thinking. This study contributes to the development of more meaningful creative thinking assessments in physics education.

Ghorbiy, B., Parno, P., & Taufiq, A. (2026). Item analysis of a creative thinking skills test on dynamic fluids material. *Momentum: Physics Education Journal*, 10(1), 27-36. <https://doi.org/10.21067/mpej.v10i1.12277>

1. Introduction

Twenty-first century education demands that students possess higher-order thinking skills, including critical, creative, collaborative, and communicative thinking. Creative thinking is an essential element in the study and practice of physics, as the field is fundamentally rooted in the discovery of innovative models and solutions to natural phenomena. Physics is a scientific endeavor that requires creativity; without it, relying solely on analysis and practice would hinder the emergence of original experiments and theories (Newton et al., 2022). Therefore, physics education must earnestly integrate opportunities for creative exploration into its curriculum and instructional practices (Doyan et al., 2024). A significant positive correlation has been found between creative thinking ability and students' learning outcomes in the topic of elasticity and Hooke's Law, where more creative students tend to demonstrate a better understanding of physics concepts (Said et al., 2024).

The topic of dynamic fluids in physics is particularly well-suited for fostering creative learning, as it presents concepts that are both abstract and highly applicable to everyday life. Examples of dynamic fluid applications include water flow in pipes, liquid sprays, the principle of airplane flight, and the working mechanism of a carburetor. These various phenomena require students to employ visualization strategies through diagrams, simulations, and imaginative reasoning in simple experiments (Misbah et al., 2023). One recent meta-analysis found that a STEM-based inquiry learning approach to dynamic fluid topics significantly enhances students' creative thinking skills (Ghorbiy et al., 2024). A study by Permana et al. (2021) showed that STEM-based inquiry helps students think creatively when designing simple tools such as pipes or nozzles. Another study by Wati et al. (2024) revealed that the use of digital worksheets in dynamic fluid instruction encouraged

students to be more creative in formulating problem-solving strategies. In addition, virtual laboratories using PhET simulations also stimulate students' creativity as they manipulate fluid parameters (Dhanil et al., 2024). Finally, Vaidya (2020) found that integrating project-based learning (PjBL) with computational simulations improved students' understanding and creativity because the simulations enabled learners to visualize invisible fluid processes, test multiple design ideas directly, and connect theoretical concepts with real-world fluid phenomena such as laminar flow and vortices.

Creative thinking in physics learning involves complex cognitive processes that enable students to generate multiple ideas, view problems from different perspectives, produce original solutions, and elaborate scientific explanations meaningfully. These processes are commonly represented through four major indicators of creative thinking: fluency, flexibility, originality, and elaboration (Tanjung & Nasution, 2023). In the context of dynamic fluid learning, these indicators are particularly important because students are required not only to apply formulas, but also to interpret fluid phenomena, design problem-solving strategies, and connect abstract concepts with real-world situations. Modern science education theories, such as constructivist and inquiry-based learning, emphasize that meaningful physics learning occurs when students actively construct knowledge through exploration, reasoning, and creative problem solving (Doyan et al., 2024; Newton et al., 2022). Therefore, assessment instruments in physics education should not merely measure procedural understanding, but also evaluate students' creative thinking processes comprehensively.

Well-designed test instruments are essential for valid and reliable measurement of students' creative thinking skills. Such tests must comprehensively reflect the indicators of creative thinking, which include fluency, flexibility, originality, and elaboration, to ensure that the results truly represent students' abilities. Tanjung & Nasution (2023) developed a creative thinking test instrument on the topic of direct current electricity and achieved a reliability coefficient of 0.756; item analysis indicated that the instrument met the criteria for validity, difficulty level, discrimination power, and was appropriate for measuring student creativity. Viyanti et al. (2022) also created a valid and reliable creativity assessment instrument for project-based physics learning, showing scores above 68% across all four creative thinking indicators. Magara et al. (2021) developed a creativity instrument for the topic of hydrocarbons, which demonstrated high expert validity ranging from 0.87–0.94. Utama et al. (2021) also produced a project-based learning instrument with a construct validity score above 0.83. In addition, the development of the ECAT (Engineering Creativity Assessment Tool) revealed four creative factors and high reliability through factor analysis and correlation with the Torrance Test of Creative Thinking (Akdemir-Beveridge et al., 2025). These studies collectively highlight that without valid and reliable instruments, students' creative thinking skills cannot be measured accurately. However, most existing creativity assessment instruments have been developed for general science or other physics topics, while instruments specifically designed and empirically analyzed for dynamic fluid learning remain limited. In classroom practice, teachers also tend to use calculation-oriented questions that emphasize procedural answers rather than students' originality and flexibility in solving physics problems. Therefore, there is an urgent need to develop and analyze essay-based instruments capable of measuring students' creative thinking skills comprehensively in the context of dynamic fluids.

Most physics test items used in schools still focus on lower-order cognitive skills, such as recalling formulas or performing calculations, without encouraging students to think originally or generate new ideas. As a result, the potential of test items to assess creative thinking has not been fully utilized. Asriadi & Istiyono (2020) noted that many physics questions have not been designed to measure student creativity, despite the fact that this skill is crucial for deep conceptual understanding. Putri et al. (2022) also found that questions used in computational physics instruction only covered a small portion of creativity indicators. Sapulete et al. (2023) reported that questions implemented in conventional physics curricula resulted in low creativity scores across all aspects. Moreover, Viyanti et al. (2022) demonstrated that developing questions within real and challenging contexts can significantly enhance students' creativity. Bhakti et al. (2024) emphasized that the content validity of many physics instruments still fails to include creative thinking indicators. This aligns with findings by Cahya et al. (2022), who showed that problem-solving instruction in physics requires questions that encourage creative, not just logical thinking. Therefore, an in-depth analysis of physics test items is essential to ensure a more accurate and comprehensive assessment of creative thinking skills.

Item analysis is essential for evaluating the extent to which physics questions are able to stimulate the indicators of creative thinking. Through item analysis, researchers and educators can identify whether each question effectively measures fluency, flexibility, originality, and elaboration, as well as determine whether the items are appropriate for distinguishing students with different levels of creative thinking ability. Without systematic item analysis, assessment results may become inaccurate and fail to represent students' actual creative potential in physics learning. A study by Viyanti et al. (2022) reported that physics questions presented in a project-based context successfully measured fluency at 69%, flexibility at 80%, originality at 69%, and elaboration at 68%. Aulia Putra & Hufri (2022) found that physics modules with high validity require item analysis to ensure that originality and elaboration are effectively measured. Based on a systematic review, Tanjung & Nasution (2023) emphasized that creative instruments must be evaluated for item alignment with the Torrance model and analyzed for difficulty level and discrimination index. Zakiyah et al. (2024) also recommended the use of context-based models to promote a balanced assessment of creative thinking dimensions through item analysis. These findings collectively suggest that without systematic item analysis, physics test questions will not fully stimulate students' flexibility and originality in thinking.

Item analysis is a crucial step in evaluating the quality of a test instrument. This process includes testing for validity, reliability, item difficulty, and discrimination power. Its main objective is to ensure that each question accurately measures the intended construct and produces trustworthy data (Paskalin & Susanti, 2020). Through this analysis, researchers can identify ambiguous or ineffective items, as well as detect imbalances in item distribution across varying levels of participant ability (Zafrullah et al., 2024). Furthermore, studies such as that by Anshari et al. (2024) emphasize the importance of item analysis for revising questions prior to broader use, in order to achieve high reliability ($\alpha \geq 0.80$) and strong empirical validity. Therefore, item analysis not only enhances the quality of a test but also strengthens research conclusions by ensuring the use of fully validated instruments.

Item analysis ensures that physics questions on dynamic fluids are feasible, clear, and objective, thus making the measurement of students' creative thinking more accurate and accountable. Empirical validity assesses the extent to which item scores correlate strongly with the total test score, indicating the item's alignment with the overall test construct (Marie & Edannur, 2015; Peligro, 2023). Reliability reflects the consistency of test results over time, where a Cronbach's alpha between 0.7–0.9 is considered adequate for internal consistency (Ashraf & K., 2020). The difficulty index indicates the proportion of students who answer a question correctly; an ideal value ranges from 0.3–0.7 to ensure the item is neither too easy nor too difficult (Reynolds et al., 2021). The discrimination index shows how well an item differentiates between high and low ability students, commonly calculated using the point-biserial index or the discrimination parameter in Item Response Theory. A value above 0.3 signifies a good-quality item (Schmucker & Moore, 2025). These four measures are standard practices within both Classical Test Theory (CTT) and Item Response Theory (IRT) to ensure the quality and fairness of items in accurately assessing student performance (Reynolds et al., 2021). By systematically applying item analysis, developers not only ensure that test items are objective and valid, but also that they effectively stimulate student creativity through the indicators of fluency, flexibility, originality, and elaboration. The purpose of this article is to evaluate the quality of test items through empirical validity, reliability, item difficulty, and discrimination indices, in order to measure students' creative thinking abilities on the topic of dynamic fluids based on the defined indicators.

2. Method

This study is a quantitative research with a descriptive approach, aimed at examining the quality of test items designed to assess students' creative thinking skills in the topic of dynamic fluids. The focus of the analysis includes empirical validity, reliability, item difficulty, and item discrimination for each test item. The instrument was developed based on creative thinking indicators, namely fluency, flexibility, originality, and elaboration, as adopted in similar studies (Akdemir-Beveridge et al., 2025).

The test items were developed through several stages, including identification of creative thinking indicators, selection of relevant dynamic fluid concepts, preparation of contextual essay questions, and development of scoring rubrics (Ramly et al., 2022; Tanjung & Nasution, 2023). The

instrument consisted of four essay items, where each item specifically represented one indicator of creative thinking, namely fluency, flexibility, originality, and elaboration. The questions were constructed using contextual dynamic fluid phenomena to encourage students to express creative ideas and problem-solving strategies. Before empirical testing, the instrument and scoring rubric were reviewed and theoretically validated by physics education experts to ensure content relevance, clarity, and alignment with the intended creative thinking constructs (Athiyah et al., 2022). Examples of the essay test items developed for each creative thinking indicator are presented in Table 1.

Table 1. Examples of Creative Thinking Test Items

Creative Thinking Indicator	Example of Test Item
Fluency (Q2)	Students are asked to propose several possible ways to create a dancing fountain with varying water heights based on the continuity principle.
Flexibility (Q4)	Students are asked to explain various tools that can be developed using the continuity principle based on experimental observations of water flow in different pipe diameters.
Originality (Q6)	Students are asked to propose original modifications to a mosquito sprayer in order to increase spray speed and range using Bernoulli's principle.
Elaboration (Q7)	Students are asked to explain in detail several causes of the decreasing shape of water flow from a faucet based on Bernoulli's principle.

The research subjects consisted of 120 students from two schools, namely MAN 2 Kota Malang and MAN Kota Batu, each contributing two classes as participants. The selection of subjects was carried out purposively, based on the criterion that the students had completed learning the topic of dynamic fluids, and were therefore considered to have the relevant prior knowledge to participate in the creative thinking skills test.

The instrument used was an essay test consisting of four items, developed based on creative thinking indicators and contextual problems related to the topic of dynamic fluids. Each item was scored on a scale of 0 to 5, according to a scoring rubric that had been developed and theoretically validated by experts prior to empirical testing (Athiyah et al., 2022; Ramly et al., 2022).

The data were obtained from students' responses to the test items and analyzed using SPSS version 30.0 to examine item validity, instrument reliability, item difficulty, and discrimination index. Item validity was tested using the Pearson Product Moment correlation between each item score and the total score. An item was considered valid if the correlation coefficient was positive and statistically significant, with a significance value (Sig. 2-tailed) < 0.05 and $r_{xy} > r_{table}$ (Wörner et al., 2022).

Instrument reliability was analyzed using Cronbach's Alpha to determine the internal consistency among the test items. The interpretation of the reliability coefficient followed these categories can be seen Table 2.

Table 2. Reliability Criteria

Cronbach's Alpha Value	Reliability Category
< 0.60	Low
$0.60-0.79$	Moderate
≥ 0.80	High

(Dorsah, 2026; Tavakol & Dennick, 2011)

Item difficulty was calculated by dividing the average score of each test item by the maximum score, which was 5. The results were classified based on the difficulty index as follows in Table 3.

Table 3. Difficulty Level Criteria

Difficulty Index	Category
$0.00-0.20$	Difficult
$0.21-0.70$	Moderate
$0.71-1.00$	Easy

(Athiyah et al., 2022)

Item discrimination was analyzed using the Corrected Item-Total Correlation values found in the SPSS output. The interpretation of discrimination index was based on the following categories in Table 4:

Table 4. Discrimination Index Criteria

Discrimination Index	Category
< 0.20	Poor
0.20–0.39	Fair
0.40–0.70	Good
≥ 0.70	Very good

(Ramly et al., 2022; Thorndike & Thorndike-Christ, 2014)

This analytical process was used to assess the appropriateness and effectiveness of each test item in comprehensively measuring students' creative thinking skills. The evaluation results also served as a basis for decision-making regarding the revision, improvement, or application of the instrument in instructional settings and further research.

3. Results and Discussion

3.1. Validity Test

The validity test aims to ensure that each item accurately reflects the construct of creative thinking as defined in the instrument design. The analysis was conducted using Pearson correlation between each item score and the total score in SPSS version 30.0. An item is considered valid if the significance value (Sig. 2-tailed) is < 0.05 and the correlation coefficient $r_{xy} > r_{table}$ (≈ 0.179 for $N = 120$). Validity test results can be seen in Table 5.

Table 5. Validity Test Results

Item Number	Pearson Correlation	r_{table}	Sig. (2-tailed)	Interpretation
Q2	0.827	0.179	< 0.001	Highly Valid
Q4	0.939			
Q6	0.946			
Q7	0.832			

All items showed a very high correlation ($r > 0.80$) with a very low significance value ($p < 0.001$). This indicates a strong and meaningful relationship between each item score and the total score, thus confirming the construct validity of the instrument empirically. The high validity may be influenced by the contextual and open-ended characteristics of the essay questions, which encouraged students to generate diverse responses and explain scientific reasoning meaningfully.

One example of the originality item (Q6) asked students to propose several modifications to a mosquito sprayer in order to increase spray speed and range based on Bernoulli's principle. This type of question required students not only to apply physics concepts, but also to produce original ideas and justify their solutions scientifically. Such characteristics enabled the instrument to capture students' creative thinking processes more comprehensively.

These validity results are supported by findings from previous studies showing that creative thinking instruments with $r > 0.80$ are generally categorized as highly valid (Kholid et al., 2022; Zain et al., 2022). Compared with previous studies, the validity results in this study were higher than those reported by Kholid et al (2022) and Zain et al (2022), which developed creative thinking assessment instruments with validity coefficients categorized as valid. The stronger validity obtained in this study may be influenced by the use of contextual and open-ended dynamic fluid problems that encouraged students to express creative ideas and scientific reasoning more comprehensively. Therefore, the instrument developed in this study demonstrates strong potential for measuring creative thinking skills in physics learning contexts.

3.2. Reliability Test

The reliability test aims to measure the internal consistency of the instrument whether the four items produce stable and interrelated scores. The analysis was conducted by calculating the

Cronbach's Alpha value using SPSS 30.0. According to Dorsah, (2026), a value of $\alpha \geq 0.80$ indicates a high and acceptable level of reliability. Reliability test results can be seen in Table 6.

Table 6. Reliability Test Results

Number of Students	Item Number	Number of Items	Cronbach's Alpha	Category
120	Q2, Q4 Q6, Q7	4	0.908	High

The α value of 0.908 indicates a very high level of internal reliability. This means that the four items consistently measured the construct of creative thinking across the indicators of fluency, flexibility, originality, and elaboration. The high reliability suggests that the contextual essay questions were capable of producing stable and interrelated responses from students during the assessment process.

One example can be seen in the flexibility item (Q4), where students were asked to propose various tools that could be developed based on the continuity principle from experimental observations of water flow in pipes with different diameters. This type of question encouraged students to explore multiple perspectives and generate diverse solutions based on the same scientific concept. As a result, students' responses showed consistent patterns related to flexibility in creative thinking.

This finding is consistent with research in STEM education and psychometrics, where $\alpha > 0.90$ reflects strong instrument cohesion (Nacion, 2024). Compared with Tanjung & Nasution (2023), who obtained a reliability coefficient of 0.756, the reliability value in this study was considerably higher. This difference may be influenced by the use of contextual dynamic fluid phenomena that are closely related to students' daily experiences, allowing students to express creative ideas more naturally and consistently.

From a pedagogical perspective, the findings indicate that contextual essay-based assessments can effectively support the evaluation of creative thinking skills in physics learning. Unlike conventional calculation-oriented questions, the instrument developed in this study encouraged students to explain ideas, propose alternative solutions, and elaborate scientific reasoning more comprehensively. Therefore, the instrument demonstrates not only strong psychometric quality but also meaningful educational value in assessing students' creative thinking skills.

3.3. Item Difficulty Level

The difficulty level analysis aims to determine whether an item is statistically considered easy or difficult. In SPSS 30.0, the difficulty index is obtained by dividing the average score of each item by the maximum score of 5, with the following classification range: difficult (≤ 0.20), moderate (0.21-0.70), and easy (0.71-1.00) (Fauzie et al., 2021). Difficulty level test results can be seen in Table 7.

Table 7. Difficulty Level Test Results

Item Number	Difficulty Index (p)	Category
Q2	0.526	Moderate
Q4	0.500	
Q6	0.480	
Q7	0.476	

The difficulty index values ranged from 0.476-0.526, indicating that all items were categorized as moderate. This result suggests that the questions were neither too easy nor too difficult for students. Moderate-level items are considered ideal because they provide students with sufficient opportunities to demonstrate their creative thinking abilities while still maintaining meaningful cognitive challenges.

The moderate difficulty level may be influenced by the contextual and open-ended nature of the essay questions, which were closely related to students' daily experiences. For example, in the elaboration item (Q7), students were asked to explain several causes of the decreasing shape of water flow from a faucet based on Bernoulli's principle. This item required students not only to recall

physics concepts, but also to elaborate scientific explanations in detail and connect multiple concepts, such as continuity and pressure differences, to real-world fluid phenomena.

Although all items were categorized as moderate, several students still experienced difficulties in providing detailed explanations and developing scientifically complete answers, particularly in the originality and elaboration indicators. Many students were able to identify basic physics concepts, but some responses lacked deeper reasoning and comprehensive explanation. This finding indicates that students are still more familiar with procedural problem solving than open-ended scientific reasoning in physics learning contexts.

These findings are consistent with Suruchi & Rana (2014), who stated that moderate difficulty items are more effective for differentiating students' abilities and supporting balanced assessment processes. Compared with conventional calculation-oriented physics questions, the instrument developed in this study provided broader opportunities for students to explore ideas, connect concepts with everyday phenomena, and express creative reasoning more meaningfully. This characteristic represents one of the main strengths of the instrument in assessing creative thinking skills in dynamic fluid learning.

3.4. Item Discrimination Index

Discrimination Index Test aims to evaluate how well each item distinguishes between students with high and low total scores. The analysis was conducted using Corrected Item-Total Correlation in SPSS 30.0. The interpretation of the correlation values is as follows: < 0.20 = poor; $0.20-0.39$ = fair; $0.40-0.70$ = good; ≥ 0.70 = very good (Reynolds et al., 2021). Discrimination index results can be seen in Table 8.

Table 8. Discrimination Index Results

Item Number	Corrected Item-Total Correlation	Interpretation
Q2	0.694	Good
Q4	0.887	Very Good
Q6	0.899	Very Good
Q7	0.703	Very Good

All correlations exceeded 0.69, with items Q4 and Q6 approaching 0.90, indicating good to very good discrimination power. These findings indicate that the essay items were effective in distinguishing students with high creative thinking abilities from those with lower levels of creative thinking. The high discrimination indices suggest that students who possessed stronger creative thinking skills were able to provide more varied, detailed, and scientifically relevant responses compared with students who tended to rely only on procedural answers.

One example can be seen in the fluency item, where students were asked to propose several possible ways to create a dancing fountain with varying water heights based on the continuity principle. Students with high creative thinking abilities were generally able to generate multiple relevant solutions, such as varying pipe diameter, using automatic nozzles, integrating sensors, and controlling water flow dynamically. In contrast, students with lower creative thinking abilities tended to provide only one simple solution with limited explanation. This difference in response quality contributed to the strong discrimination power of the item.

These findings are consistent with previous studies showing that open-ended and contextual physics questions tend to produce higher discrimination indices because they allow clearer differentiation between students' thinking abilities (Dalyanto et al., 2021; R. Permana, 2023). Compared with conventional physics questions that mainly emphasize formula application, the instrument developed in this study encouraged students to explore multiple ideas, elaborate scientific reasoning, and connect physics concepts with real-world situations more meaningfully.

From a pedagogical perspective, the high discrimination indices indicate that the instrument can help teachers identify variations in students' creative thinking abilities more accurately. This is important because creative thinking assessment in physics learning should not only focus on

obtaining correct answers, but also evaluate how students generate ideas, develop explanations, and propose alternative solutions scientifically.

The novelty of this instrument lies in its integration of contextual dynamic fluid phenomena with specific creative thinking indicators in essay-based assessment. Each item was intentionally designed to represent one indicator of creative thinking, namely fluency, flexibility, originality, and elaboration, enabling more focused evaluation of students' creative thinking processes in dynamic fluid learning contexts.

4. Conclusion

This study examined the quality of essay test items designed to assess students' creative thinking skills in dynamic fluid learning through validity, reliability, difficulty level, and discrimination index analyses. The developed instrument consisted of four essay questions representing the indicators of fluency, flexibility, originality, and elaboration, and was tested on 120 students who had studied dynamic fluids. The results showed that all items had very high validity coefficients ($r = 0.827-0.946$; $p < 0.001$), a very high reliability coefficient (Cronbach's Alpha = 0.908), moderate difficulty levels (0.476–0.526), and good to very good discrimination indices (0.694–0.899). These findings indicate that the instrument is valid, reliable, and effective for measuring students' creative thinking skills in physics learning. The novelty of this study lies in the development of a contextual essay-based assessment instrument specifically designed for dynamic fluid topics, where each item explicitly represents one indicator of creative thinking and encourages students to generate diverse ideas, elaborate scientific reasoning, and connect physics concepts with real-life phenomena. However, this study was limited to two schools and only focused on dynamic fluid material with four essay items, so the instrument may not yet represent broader dimensions of creative thinking across different educational contexts. Therefore, future studies are recommended to involve more diverse participants, develop additional essay items, and apply the instrument to other physics topics or technology-based assessment environments to obtain more comprehensive evaluation results of students' creative thinking skills.

Author Contributions

All authors have equal contributions to the paper. All the authors have read and approved the final manuscript.

Funding

No funding support was received.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgement

The author would like to express sincere gratitude to the academic advisor for their guidance and direction in the writing of this article. Special thanks are also extended to the author's sibling for their assistance in drafting the introduction, searching for scientific article references, and thoroughly reviewing the manuscript before submission. Heartfelt appreciation is also given to the author's parents and family for their moral support, encouragement, and prayers throughout the writing process of this article.

References

- Akdemir-Beveridge, Z. G., Zaghi, A., & Syharat, C. (2025). Understanding and Evaluating Engineering Creativity: Development and Validation of the Engineering Creativity Assessment Tool (ECAT). *arXiv preprint arXiv: https://doi.org/10.48550/arXiv.2504.12481*
- Anshari, M. I., Nasution, R., Irsyad, M., Alifa, A. Z., & Zuhriyah, I. A. (2024). Analisis Validitas dan Reliabilitas Butir Soal Sumatif Akhir Semester Ganjil Mata Pelajaran PAI. *Edukatif: Jurnal Ilmu Pendidikan*, 6(1), 964–975. <https://doi.org/10.31004/edukatif.v6i1.5931>
- Ashraf, Z. A., & K., J. (2020). Classical and Modern Methods in Item Analysis of Test Tools Assistant Professor of Clinical Psychology, IMHANS, Kozhikode. *International Journal of Research and Review (Ijrrjournal.Com)*, 7(5), 397–403.

- Asriadi, M., & Istiyono, E. (2020). Exploration of Creative Thinking Skills of Students in Physics Learning. *Journal of Educational Science and Technology (EST)*, 6, 151–158. <https://doi.org/10.26858/est.v6i2.12737>
- Athiyah, R., Feranie, S., & Ramalis, T. R. (2022). Development of Scientific Process-Creative Skills (SP-CS) Test on Light Wave Concept: Content Validity and Rasch Model Analysis. *Edusains*, 14(2), 111–125. <https://doi.org/10.15408/es.v13i2.28025>
- Bhakti, Y. B., Arthur, R., & Supriyati, Y. (2024). The Content Validity: Multiple Choices Complex Physics Instrument to Measure Critical Thinking Skills. *JIPF (Jurnal Ilmu Pendidikan Fisika)*, 9(1), 56–64. <https://doi.org/10.26737/jipf.v9i1.4765>
- Cahya, R., Rokhmat, J., & Gunada, I. W. (2022). Validity of Learning Tools Creative Problem Solving Models to Improve Students' Physics Problem-Solving Ability. *Jurnal Pendidikan Fisika*, 10(1), 43–48. <https://doi.org/10.24252/jpf.v10i1.27246>
- Dhanil, M., Putri, F. A., Desnita, D., & Usmeldi, U. (2024). Analysis of the Use of Dynamic Fluid Learning Media. *Physics Education Research Journal*, 6(1), 39–46. <https://doi.org/10.21580/perj.2024.6.1.16364>
- Dorsah, P. (2026). The Use of Cronbach's Alpha Reliability in Educational Research: A Systematic Review. *European Journal of Contemporary Education and E-Learning*, 4(2), 39–50. [https://doi.org/10.59324/ejceel.2026.4\(2\).04](https://doi.org/10.59324/ejceel.2026.4(2).04)
- Doyan, A., Susilawati, S., Annam, S., Mulyadi, L., Megahati, R. R. P., Hutabarat, R. A., Ikhsan, M., Ardianti, N. R., & Hamidi, H. (2024). Trends Research Creative Thinking Skills in Students' Physics Learning (2015-2024): A Systematic Review. *Jurnal Penelitian Pendidikan IPA*, 10(6), 333–342. <https://doi.org/10.29303/jppipa.v10i6.7826>
- Fauzie, M., Pada, A. U. T., & Supriatno, S. (2021). Analysis of the Difficulty Index of Item Bnk According to Cognitive Aspects during the Covid-19 Pndemic. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 25(2), 162–169. <https://doi.org/10.21831/pep.v25i2.42603>
- Ghorbiy, B., Sutopo, S., & Purwaningsih, E. (2024). Systematic Literature Review: Analysis of Implementation Trends of STEM-Based Physics Learning on Dynamic Fluid Material. *Jurnal Riset Dan Kajian Pendidikan Fisika*, 11(2), 71–79. <https://doi.org/10.12928/jrpkpf.v11i2.809>
- Kholid, K., Hidayat, S., & Syafrizal, S. (2022). Pengembangan Instrumen Penilaian Kreativitas B. *Eduscience: Jurnal Ilmu Pendidikan*, 8(1). <https://doi.org/10.47007/edu.v8i02.6907>
- Magara, E., Copriady, J., & Linda, R. (2021). Validity of Assessment Instruments for Students Creative Thinking Ability on Hydrocarbon Material. *Al-Ishlah: Jurnal Pendidikan*, 13(1), 49–60. <https://doi.org/10.35445/alishlah.v13i1.264>
- Marie, S. M. J. A., & Edannur, S. (2015). Relevance of Item Analysis in Standardizing an Achievement Test in Teaching of Physical Science in B.Ed Syllabus. *I-Manager's Journal of Educational Technology*, 12(3), 30–36. <https://doi.org/10.26634/jet.12.3.3743>
- Misbah, M., Hamidah, I., Sriyati, S., & Samsudin, A. (2023). Research Trend of Dynamic Fluid in Learning: A Bibliometric Analysis. *JPPPF (Jurnal Penelitian Dan Pengembangan Pendidikan Fisika)*, 9(2), 263–272. <https://doi.org/10.21009/1>
- Nacion, R. H. (2024). Development and Validation of an Assessment Tool for Measuring Innovative Thinking in STEM Students. *Applied Research and Innovation*, 2(2), 1–12. <https://doi.org/10.54536/ari.v2i2.3471>
- Newton, D. P., Nolan, S., & Rees, S. (2022). Creative thinking in physics. In *Creative Thinking in University Physics Education* (pp. 1-1-1-13). *IOP Publishing*. <https://doi.org/10.1088/978-0-7503-4028-1ch1>
- Paskalin, G., & Susanti, M. M. I. (2020). Item Analysis of Force Material Problem in Elementary School. *DIDAKTIKA: Jurnal Pendidikan Sekolah Dasar*, 3(1), 23–32. <https://doi.org/10.21831/didaktika.v3i1.31431>
- Peligro, V. C. (2023). The Role of Item Analysis in Detecting and Improving Faulty Multiple-Choice Test. *Teaching and Teacher Education*, (132). <https://www.sciencedirect.com/journal/social-science-and-medicine>
- Permana, G. A., Parno, P., Hidayat, A., & Ali, M. (2021). Improving Creative Thinking Skill of Fluid Dynamic through IBL-STEM with Formative Assessment. *AIP Conference Proceedings*, 2330, 050016. <https://doi.org/10.1063/5.0043128>
- Putra, Y. A., & Hufri, H. (2022). Validity of Physics Module Integrating Creative Thinking Ability in Momentum and Impulse Materials. *Pillar of Physics Education*, 15(3), 200–208. <https://doi.org/10.24036/13515171074>
- Putri, R. M., Akmam, A., Mufit, F., Sari, S. Y., & Hidayat, R. (2022). Analysis of Student's Creative Thinking Ability in Computation Physics Course. *Gravity: Jurnal Ilmiah Penelitian Dan Pembelajaran Fisika*, 8(2), 78–86. <https://doi.org/10.30870/gravity.v8i2.13783>
- Ramly, S. N. F., Ahmad, N. J., & Yakob, N. (2022). Development, Validity, and Reliability of Chemistry Scientific Creativity Test for Pre-University Students. *International Journal of Science Education*, 44(14), 1–16. <https://doi.org/10.1080/09500693.2022.2116298>
- Reynolds, C. R., Altmann, R. A., & Allen, D. N. (2021). Item Analysis: Methods for Fitting the Right Items to the Right Test. In C. R. Reynolds, R. A. Altmann, & D. N. Allen (Eds.), *Mastering Modern Psychological Testing* (pp. 263–289). *Springer International Publishing*. https://doi.org/10.1007/978-3-030-59455-8_7
- Said, M. A., Ilham, I., & Kadir, M. F. A. (2024). Fostering Creativity to Enhance Physics Achievement: An Analysis of the Relationship Between Creative Thinking Ability and Student Learning Outcomes. *Unnes Science Education Journal*, 13(3), 180–188. <https://doi.org/10.15294/usej.v13i3.14516>
- Sapulete, H., Wenno, I. H., Tuhurima, D., & Dulhasyim, A. B. P. (2023). Analysis of Creative Thinking Skills Based on Contextual Teaching and Learning in Physics Education Students. *Jurnal Penelitian Pendidikan IPA*, 9(7), 5492–5497. <https://doi.org/10.29303/jppipa.v9i7.3691>

- Schmucker, R., & Moore, S. (2025). The Impact of Item-Writing Flaws on Difficulty and Discrimination in Item Response Theory. *ArXiv*. <https://doi.org/https://doi.org/10.48550/arXiv.2503.10533>
- Suruchi, S., & Rana, S. S. (2014). Test Item Analysis and Relationship Between Difficulty Level and Discrimination Index of Test Items in an Achievement Test in Biology Education. *Paripex - Indian Journal of Research*, 3(6), 56–58.
- Tanjung, Y. I., & Nasution, I. R. (2023). The Development of Creative Thinking Test Instruments with Torrance Indicators on Direct Current Electricity Materials. *Jurnal Pendidikan Fisika Indonesia*, 18(2), 134–143. <https://doi.org/10.15294/jpfi.v18i2.32117>
- Tavakol, M., & Dennick, R. (2011). Making Sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Thorndike, R. M., & Thorndike-Christ, Tracy. (2014). *Measurement and Evaluation in Psychology and Education* (8th ed.). Pearson.
- Utama, A. S., Rosidin, U., & Suyatna, A. (2021). Instrument of Creative Thinking Skills in Project Based Learning Physics. *Indonesian Journal of Science and Mathematics Education*, 4(1), 58–68. <https://doi.org/10.24042/ij sme.v4i1.8572>
- Vaidya, A. (2020). Teaching and Learning of Fluid Mechanics. *Fluids*, 5(2), 49. <https://doi.org/10.3390/fluids5020049>
- Viyanti, Herliani, D., & Rosidin, U. (2022). Strengthening Creative Thinking Ability: A Project-Based Physics Assessment Instrument. *Impulse: Journal of Research and Innovation in Physics Education*, 2(2), 107–118. <https://doi.org/10.14421/impulse.2022.22.05>
- Wati, M., Lestari, D. M., & Mahtari, S. (2024). The Effect of Digital Worksheets on Dynamic Fluids with Authentic Learning in Training Problem-Solving Skills. *Jurnal Ilmiah Pendidikan Fisika*, 8(2), 266. <https://doi.org/10.20527/jipf.v8i2.9245>
- Wörner, S., Becker, S., Küchemann, S., Scheiter, K., & Kuhn, J. (2022). Development and validation of the Converging Lenses Concept Inventory for Middle School Physics Education. *ArXiv*. <https://doi.org/10.48550/arXiv.2205.09861>
- Zafrullah, Z., Kastara, R., Riantoro, E. S., & Bakti, A. A. (2024). Analisis Butir Soal dengan ITEMAN 4.0 pada Penilaian Akhir Semester Sekolah Dasar. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, 9(1), 5813–5823. <https://doi.org/10.23969/jp.v9i1.8777>
- Zain, A. Z., Ramalis, T. R., & Muslim, M. (2022). Karakterisasi Instrumen Tes Keterampilan Bepikir Kreatif Berdasarkan Analisis Partial Credit Model. *Jurnal Ilmiah Pendidikan Fisika*, 6(1), 176. <https://doi.org/10.20527/jipf.v6i1.4806>
- Zakiah, S., Suwarma, I. R., & Rusdiana, D. (2024). Constructing Assessment for Evaluating Critical Thinking and Creative Problem-Solving Skills in the Lesson on Soundwave. *Jurnal Penelitian Pendidikan IPA*, 10(8), 4614–4622. <https://doi.org/10.29303/jppipa.v10i8.8332>