**Kasuari: Physics Education Journal (KPEJ) 9(2) (2026) 317-326**
P-ISSN: 2615-2681
E-ISSN: 2615-2673

**Kasuari: Physics Education Journal (KPEJ)**
**Universitas Papua**
website: https://journalfkipunipa.org/index.php/kpej

# Empirical Analysis of the Validity and Reliability of Creative Thinking Skills Test Instruments on Static Fluid Material for High School Students

**Dya Atul Fahmi[1], Parno[1*], Khusaini[1], & Marlina Ali[2]**
[1]Departemen of Physics, Universitas Negeri Malang, Indonesia
[2]School of Education, University Teknologi Malaysia, Malaysia
*Corresponding author: parno.fmipa@um.ac.id

**Abstract:** The accuracy of learning assessment results is greatly influenced by the quality and reliability of the instruments used, so instrument validation is a crucial step in measuring students' creative thinking abilities. However, many instruments have not been properly validated, potentially producing inaccurate data and biased conclusions. This study aims to analyze the validity and reliability of a creative thinking ability test instrument about static fluids for high school students using a quantitative descriptive approach. The instrument consists of eight fill-in-the-blank questions structured based on creative thinking indicators and completed independently by students. Data analysis included item validity testing, reliability testing, difficulty level analysis, and discriminatory power analysis. The results show that all items are valid, the instrument is reliable, most items are at a moderate level of difficulty, and have adequate to very high discriminatory power. Thus, this test instrument is suitable for use as a tool for assessing high school students' creative thinking abilities.

**Keywords:** creative thinking skills, high school students, reliability, static fluid, test instrument, validity

# Analisis Empiris Validitas dan Reliabilitas Instrumen Tes Keterampilan Berpikir Kreatif pada Material Fluida Statis untuk Siswa SMA

***Abstrak:*** *Akurasi hasil penilaian pembelajaran sangat dipengaruhi oleh kualitas dan reliabilitas instrumen yang digunakan, sehingga validasi instrumen merupakan langkah penting dalam mengukur kemampuan berpikir kreatif siswa. Namun, banyak instrumen belum divalidasi dengan benar, berpotensi menghasilkan data yang tidak akurat dan kesimpulan yang bias. Studi ini memiliki tujuan untuk menganalisis validitas dari instrumen tes kemampuan berpikir kreatif tentang fluida statis untuk siswa SMA menggunakan pendekatan deskriptif kuantitatif. Instrumen tersebut terdiri dari delapan pertanyaan isian yang disusun berdasarkan indikator berpikir kreatif dan dikerjakan secara mandiri oleh siswa. Analisis data meliputi pengujian validitas item, pengujian reliabilitas, analisis tingkat kesulitan, dan analisis daya diskriminatif. Hasil menunjukkan bahwa semua item valid, instrumen reliabel, sebagian besar item berada pada tingkat kesulitan sedang, dan memiliki daya diskriminatif yang memadai hingga sangat tinggi. Dengan demikian, instrumen tes ini cocok digunakan sebagai alat untuk menilai kemampuan berpikir kreatif siswa SMA.*

***Keywords:*** *fluida statis, instrument tes, keterampilan berpikir kreatif, reliabilitas, siswa SMA, validitas*

## INTRODUCTION

Creative thinking skills are high-level cognitive skills that are crucial in 21st-century education because they play a role in generating new ideas, fostering flexibility in thinking, and fostering innovative problem-solving abilities (Burakgazi & Reiss, 2025; Ramos & Araya, 2025; Suratno et al., 2019). In physics learning, creative thinking skills help students connect abstract concepts with real-world phenomena, so that learning does not stop at memorizing formulas. Static fluid material, which includes hydrostatic pressure, Pascal's law, and Archimedes' principle, requires students to connect concepts to various everyday contexts, such as hydraulic systems, dams, and floating-sinking phenomena. Therefore, understanding static fluid material requires creative thinking skills so that students can reason concepts flexibly and meaningfully (Firdaus, 2025). In determining how difficult a student is, an instrument is needed that can be used as a measuring tool.

To accurately measure students' creative thinking skills, high-quality assessment instruments are required (Barbot & Kaufman, 2025; Suherman & Vidakovich, 2024). Creative thinking ability test instruments must accurately represent the construct being measured and include key indicators such as fluency, flexibility, originality, and elaboration (Haq & Pifarré, 2023). The quality of an instrument is determined by its validity and reliability, as these two aspects ensure that the measurement results truly reflect students' abilities and are consistent when used across settings (Camacho-Tamayo & Bernal-Ballen, 2023). Instruments that do not meet validity and reliability criteria have the potential to produce biased assessments and cannot be used as a basis for learning decisions (Ramos & Araya, 2025; Subhaktiyasa, 2024).

However, various studies have shown that many creative thinking ability test instruments used in science and physics learning have not undergone adequate empirical validation (Barbot & Kaufman, 2025). Some instruments are validated solely through expert assessment without the support of statistical analysis based on student response data (Mura et al., 2025; Tumangger et al., 2025). Other studies also report low reliability of creative thinking ability instruments, making them unable to accurately distinguish student ability levels (Hardila & Vebrianto, 2024; Magara et al., 2021). This condition results in inaccurate assessment results and ineffective learning feedback provided to students.

A possible solution to address these issues is empirical validation of creative thinking skills test instruments. Empirical validation aims to ensure the technical quality of the instrument through appropriate statistical testing (Allo et al., 2025; Firdaus, 2025). Through this process, items that do not align with the construct can be revised or eliminated, resulting in a more accurate and stable instrument (Camacho-Tamayo & Bernal-Ballen, 2023). Empirically validated instruments can also serve as diagnostic tools to identify students' creative thinking ability profiles and support evidence-based assessment in physics learning.

Based on the above description, this study aims to empirically analyze the validity and reliability of a creative thinking skills test instrument on static fluids for high school students. This study is expected to produce a valid, reliable, and suitable test instrument for use in physics learning. Furthermore, the results are expected to serve as a reference for teachers and researchers in developing more accurate learning assessments and supporting the sustainable development of students' creative thinking skills (Firdaus, 2025).

## METHOD

This study used a quantitative descriptive approach to evaluate and validate a creative thinking test instrument about static fluids for high school students. This approach was chosen because it is suitable for analyzing the instrument's characteristics empirically based on quantitative measurement data. The main focus of this study is to assess the validity and quality of the test instrument as a measure of students' creative thinking abilities in the context of physics learning. The instrument developed consisted of eight fill in the blank questions, designed based on creative thinking indicators and tailored to the competencies of static fluids material. Details of the indicators for each question item are presented in Table 1.

**Table 1.** Creative Thinking Skills Test Indicators

| Indicator | Item Number |
|---|---|
| Presented with the problem of landslides that cause people's houses to collapse, students create technological designs to solve this problem | 1 |
| Presented with the teapot effect phenomenon, students provided various different explanations using different approaches for each explanation category | 2 |
| Students were presented with the phenomenon of a dam that already had a pre-existing function but could not yet be used for irrigation. Students detailed the dam's components to fulfill this new function | 3 |
| A picture of a hydraulic lift was presented. Students analyzed Pascal's law in solving everyday physics problems | 4 |
| The phenomenon of someone pressing the center of a plastic bottle filled with water was presented, causing water to flow out of a small hole in the bottle cap. Students provided various explanations for the phenomenon of pressure in a fluid in a closed system | 5 |
| A tool and materials for building an excavator were presented. Students described and provided ideas | 6 |
| A case study was presented of a pontoon bridge that was prone to tipping over when vehicles passed over it. Students provided detailed explanations for improving the pontoon bridge design by adding technical details to several aspects | 7 |
| The phenomenon of a ping-pong ball being pressed into water and then released immediately floats to the surface. Students provided various ideas about what causes objects to float | 8 |

Empirical validation of the instrument was conducted on 70 high school students in two schools: one in Malang City and one in Blitar Regency, who had received static fluid instruction. Prior to the empirical test, the instrument underwent expert validation by lecturers on December 1 to ensure the appropriateness of the content, construction, and language of the questions. Subsequently, empirical testing was conducted on students on December 5 using a closed-ended system to ensure the results reflected students' independent creative thinking abilities. The empirical validation process refers to four main stages: item validity testing, instrument reliability testing, discrimination testing, and difficulty level testing. These four stages are standard procedures often used in instrument evaluation research to ensure the instrument is suitable for use in measuring higher-order cognitive abilities (Crocker & Algina, 1986).

The first stage is the item validity test. This stage aims to ensure each item is able to measure aspects of creative thinking according to the static fluid material indicators. Validity indicates the relationship between the item score and the total instrument score, thus ensuring reliable measurement results. Validity criteria are determined based on the correlation coefficient (r) value. An item is declared valid if the calculated r is greater than the table r, and invalid otherwise.

The second stage is the instrument reliability test, which aims to determine the instrument's consistency as a measuring tool. Reliability indicates the stability of the instrument's scores in measuring the construct of creative thinking. Reliability criteria were determined using the Cronbach's Alpha coefficient in Table 2.

**Table 2.** Reliability Test Criteria

| Cronbach Alpha | Category |
|---|---|
| $0.00 \leq r < 0.20$ | Unreliable |
| $0.20 \leq r < 0.40$ | Less reliable |
| $0.40 \leq r < 0.60$ | Fairly reliable |
| $0.60 \leq r < 0.80$ | Reliable |
| $0.80 \leq r \leq 1.00$ | Very reliable |

The third stage is the difficulty level test, which aims to determine the level of ease or difficulty of each item. This analysis ensures that the questions are appropriate for high school students' abilities and are neither too easy nor too difficult. The difficulty level criteria are presented in Table 3. All analyses were conducted using IBM SPSS 27.

**Table 3.** Test Criteria for Difficulty Level

| Difficulty Level | Category |
|---|---|
| $0.00 \leq D \leq 0.30$ | Difficult |
| $0.30 < D \leq 0.70$ | Medium |
| $0.70 < D \leq 1.00$ | Easy |

The fourt stage is the item discrimination test, which aims to determine the question's ability to differentiate between students with high and low levels of creative thinking. The discrimination test indicates the item's effectiveness in identifying differences in student ability. The discrimination criteria are presented in Table 4.

**Table 4.** Discrimination Index Criteria

| Discrimination Index | Category |
|---|---|
| $D \leq 0{,}00$ | Very low |
| $0.00 < D \leq 0.20$ | Low |
| $0.20 < D \leq 0.40$ | Medium |
| $0.40 < D \leq 0.70$ | High |
| $0.70 < D \leq 1.00$ | Very high |

## RESULTS AND DISCUSSION

The results of this study are presented based on four stages of instrument analysis: item validity testing, instrument reliability testing, difficulty level testing, and discrimination index testing. These four stages were conducted to ensure that the creative thinking test instrument on static fluids was suitable for use as a measure of student ability. The results of the item validity testing are presented in Table 5.

**Table 5.** Empirical Validity Test Result

| Item Number | r value | Category |
|---|---|---|
| 1 | 0.498 | Valid |
| 2 | 0.590 | Valid |
| 3 | 0.628 | Valid |
| 4 | 0.735 | Valid |
| 5 | 0.588 | Valid |
| 6 | 0.731 | Valid |
| 7 | 0.762 | Valid |
| 8 | 0.645 | Valid |

Based on Table 5, item validity testing was conducted using Pearson correlation between the scores for each item and the total score. Based on 70 respondents, the table r value was 0.235 ($\alpha = 0.05$). The analysis results show that all items had calculated r values greater than the table r, indicating that all items are empirically valid. This indicates that each item is able to measure the intended aspect of creative thinking and aligns with the static fluid learning indicators, making it suitable for use in measuring students' creative thinking abilities.

**Table 6.** Reliability Test Result

| Cronbach's Alpha | Number of Items | Category |
|---|---|---|
| 0.797 | 8 | Reliable |

Next, the results of the instrument reliability test are presented in Table 6. Reliability testing was conducted using the Cronbach's Alpha coefficient to determine the instrument's level of internal consistency. This process involved analyzing the interrelationships between items in measuring the construct of creative thinking. The analysis showed a Cronbach's Alpha value of 0.797, which is considered reliable. Therefore, the instrument has good consistency as a measurement tool. Therefore, the creative thinking test instrument on static fluids can produce stable and reliable data.

**Table 7.** Item Difficulty Index

| Item Number | Difficulty Index | Category |
|---|---|---|
| 1 | 0.79 | Easy |
| 2 | 0.35 | Medium |
| 3 | 0.59 | Medium |
| 4 | 0.60 | Medium |
| 5 | 0.67 | Medium |
| 6 | 0.58 | Medium |
| 7 | 0.54 | Medium |
| 8 | 0.58 | Medium |

The results of the item difficulty test are presented in Table 7. The difficulty analysis was conducted to determine the proportion of students who were able to answer each item correctly. Based on the analysis, one item, question number 1, was classified as easy, while the other seven items fell into the medium category. These results indicate that the instrument's overall difficulty level met standards, as most items fell into the medium category, considered ideal for optimally measuring students' creative thinking abilities.

**Table 8.** Item Discrimination Index

| Item Number | Corrected Item-Total Correlation | Category |
|---|---|---|
| 1 | 0.336 | Medium |
| 2 | 0.425 | Medium |
| 3 | 0.514 | High |
| 4 | 0.617 | Very high |
| 5 | 0.445 | Medium |
| 6 | 0.592 | High |
| 7 | 0.610 | Very high |
| 8 | 0.528 | High |

The results of the discrimination index test are presented in Table 8, which were analyzed using the Corrected Item-Total Correlation value. Based on the test results, the eight items had adequate to very high discrimination power, with coefficient values ranging from 0.336 to 0.617. No items were in the low or very low category, so all items were declared capable of distinguishing students with high and low abilities well. Therefore, all items were retained and suitable for use in further analysis.

**Discussion**

The results of the study indicate that the creative thinking ability test instrument for static fluids, consisting of eight short-answer questions, meets the appropriate validity criteria for use as an assessment instrument. Students completed the instrument independently, so the scores obtained reflect their creative thinking abilities without any assistance. This indicates that the overall construction of the instrument aligns with the concepts being measured and is able to illustrate differences in creative thinking abilities among students. When compared with previous research, these findings not only confirm the instrument's general feasibility but also demonstrate its more specific contribution to the static fluid context. Several previous studies have shown that measuring creative thinking abilities in physics learning is often conducted in a general context or focuses on learning models without a strong connection to specific material, potentially failing to fully capture the contextual nature of the creative thinking process (Astalini et al., 2022; Sabila et al., 2024).

In this study, the instrument was designed by integrating static fluid concepts such as hydrostatic pressure, Pascal's law, and Archimedes' law, which require visualization skills, flexibility of thinking, and elaboration of ideas. This approach aligns with findings that measuring creative thinking skills becomes more representative when linked to specific knowledge domains and physics learning contexts (Kurniawan et al., 2025; Sari et al., 2024). Therefore, the novelty of this research lies in the development of an instrument that is not only psychometrically valid but also has stronger contextual relevance in physics learning.

The item validity test in this study was conducted using the Pearson Product Moment correlation between the score of each item and the total score. The analysis results showed that all items were statistically valid because the correlation coefficient for each item was greater than the table r. This finding is in line with research stating that item validity analysis can be conducted through the relationship between item scores and the total score (item-total correlation), which indicates the extent to which an item represents the construct being measured (Marianti et al., 2023). In the context of learning evaluation, the use of Pearson correlation is also commonly used to determine item validity, where an item is declared valid if the correlation value obtained meets predetermined criteria (Lusiana et al., 2024). However, this study still focused on internal validity based on statistical relationships between scores, so it did not in-depth examine the relationship between the item and the context of the material or the construct being measured. Therefore, in this study, validity is not only defined as a statistical relationship but also as a representation of the suitability between the item and the construct of creative thinking in the context of statistical fluidity.

The reliability of the creative thinking test instrument in this study ranged from good to excellent, indicating internal consistency across items measuring the same construct. A high Cronbach's alpha value is widely used as an indicator of instrument consistency in educational research, particularly in demonstrating internal consistency across items (Setyaedhi, 2024). However, research in education shows that instrument reliability is also influenced by item characteristics and variations in student responses. Therefore, open-ended instruments require clear scoring guidelines to maintain consistency in measurement results (Zakariya, 2022). Therefore, the high reliability achieved in this study indicates that the instrument and rubric developed are able to maintain assessment consistency and support the use of short-answer questions in statistical fluidity materials to measure creative thinking more authentically.

Analysis of item difficulty levels showed that most items fell into the moderate category, with only one item (number 1) classified as easy and seven others at a moderate level. This finding aligns with the principles of educational evaluation, which state that a moderate level of difficulty provides optimal information for distinguishing student abilities. However, in the context of measuring creative thinking skills, difficulty is determined not only by cognitive complexity but also by the demand for diverse, flexible, and original ideas. This contrasts with previous research, which suggests that creative thinking tasks tend to be difficult for students due to difficulties in developing and expressing ideas independently (Sari et al., 2023) and are influenced by assessment designs that are not aligned with student characteristics (Widyastuti & Jayanta, 2024). The weaknesses in these studies were successfully addressed in this study by designing items that were more contextual and tailored to the characteristics of high school students, ensuring that the items remained challenging but did not create excessive cognitive load.

The eight items in this study demonstrated moderate to very high discriminatory power, meaning each item was able to effectively differentiate students with high and low creative thinking abilities. This good discrimination index indicates that the test items have high sensitivity in identifying variations in student ability. However, this finding does not always align with previous research. Previous research has shown that measuring creative thinking ability often faces challenges in consistent discrimination, particularly with open-ended instruments that generate a wide and complex range of answers, complicating the assessment process (Gyamfi & Wren, 2022). Furthermore, developing creative thinking instruments requires systematic and structured design to maintain validity and reliability, as without good instrument design, measurement tends to be inconsistent (Viyanti et al.,

2022). Furthermore, in the context of physics education, the development of more appropriate instruments to accurately and representatively measure students' creative thinking ability is still needed (Susilowati et al., 2022). These weaknesses were successfully minimized in this study through systematic question design, including clarity of creative thinking indicators, appropriate context for statistical variables, and focused yet open-ended question structures, resulting in strong and consistent psychometric qualities in differentiating student abilities.

The implications of this research are quite important for physics learning in secondary schools. This instrument can be used as an alternative assessment to measure creative thinking skills, which have so far been measured systematically. Furthermore, the use of questions based on the context of fluid statistics can encourage students to not only memorize formulas but also understand the concepts in depth and apply them to new situations. Furthermore, this research also opens up opportunities to develop similar instruments for other physics materials, thus supporting the development of 21st-century skills-based assessments more broadly. However, this study still has limitations, particularly the limited number of items and the material coverage that only covers fluid statistics. Therefore, further research is recommended to expand the material coverage and number of items, as well as test the instrument on a more diverse sample. Overall, the novelty of this research lies in the integration between the context of the fluid statistics material, the question format, and the creative thinking indicators, resulting in an instrument that is not only valid and reliable but also relevant to the needs of assessment in secondary schools.

## CONCLUSION AND SUGGESTIONS

Based on the results of the research that has been conducted, it is concluded that the instrument for testing the creative thinking abilities of high school students for static fluid material has met the eligibility criteria as an assessment tool. All items have been proven valid and the instrument as a whole shows good reliability. In addition, the analysis of the difficulty level shows that most items are in the medium category, and all items have adequate to very high discrimination power. Thus, this test instrument is suitable for measuring students' creative thinking abilities objectively and consistently. Recommendations for further research are that the instrument be tested on a larger number of respondents and involve students with more diverse backgrounds and school characteristics, so that the measurement results obtained have a broader and more accurate level of generalization.

## BIBLIOGRAPHY

Allo, A. Y. T., Pramanagara, R. O., & Suhendra, C. D. (2025). Development of Augmented Reality Media to Enhance Students' Critical Thinking Skills in High Schools. *Kasuari: Physics Education Journal (KPEJ)*, *8*(1), 40–49. https://doi.org/10.37891/kpej.v8i1.494

Astalini, Darmaji, Kurniawan, D. A., & Minarsih. (2022). Identification of HOTS Creative Thinking , Science Process Skills and Digital Literacy in Physics Subject. *Jurnal Penelitian Fisika dan Aplikasinya (JPFA)*, *12*(1), 47–61. https://doi.org/10.26740/jpfa.v12n1.p47

Barbot, B., & Kaufman, J. C. (2025). PISA 2022 Creative Thinking Assessment: Opportunities, Challenges, and Cautions. *The Journal of Creative Behavior*, *59*(1), 1–8. https://doi.org/10.1002/jocb.70003

Burakgazi, S. G., & Reiss, M. J. (2025). Exploring Creative Thinking Skills in PISA : an

Ecological Perspective on High-Performing Countries. *Frontiers in Psychology*, 1–12. https://doi.org/10.3389/fpsyg.2025.1554654

Camacho-Tamayo, E., & Bernal-Ballen, A. (2023). Validation of an Instrument to Measure Natural Science Teachers ' Self-Perception about Implementing STEAM Approach in Pedagogical Practices. *Education Sciences*, *13*(764), 1–16. https://doi.org/https://doi.org/10.3390/educsci13080764

Crocker, L., & Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. Holt, Rinehart and Winston.

Firdaus, F. (2025). Design of Instruments for Scientific Creative Thinking Skills and Creative Thinking Digital Skills : Rasch Models and Confirmatory Factor Analysis. *EURASIA Journal of Mathematics, Science and Technology Education*, *21*(5), 1–16. https://doi.org/https://doi.org/10.29333/ejmste/16310

Gyamfi, A., & Wren, D. G. (2022). Determining the Difficulty and Discrimination Parameters of a Mathematics Performance-Based Assessment. *Creative Education*, 3483–3489. https://doi.org/10.4236/ce.2022.1311223

Haq, I. U., & Pifarré, M. (2023). Dynamics of Automatized Measures of Creativity : Mapping the Landscape to Quantify Creative Ideation. *Frontiers in Education*, 1–16. https://doi.org/10.3389/feduc.2023.1240962

Hardila, D., & Vebrianto, R. (2024). Validity of Student Creativity Test Instruments on Energy and its Changes for Junior Sigh school Students. *Jurnal Mangifera Edu*, *9*(1), 56–63. https://doi.org/10.31943/mangiferaedu

Kurniawan, Sinaga, P., & Saepuzaman, D. (2025). Aiken's V Analysis and Rasch Modeling to Determine the Quality of Physics Creative Thinking Test Instruments. *Jurnal Riset dan Kajian Pendidikan Fisika*, *12*(1), 32–40. https://doi.org/10.12928/jrkpf.v12i1.1161

Lusiana, I. S., Susongko, P., & Nafiati, D. A. (2024). Desain Instrumen Tes Capaian Pembelajaran Matematika Dengan Uji Validitas Pearson Correlation. *Journal of Education Research*, *5*(4), 5666–5675.

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

Marianti, S., Rufaida, A., Hasanah, N., & Nuryanti, S. (2023). Comparing Item-total Correlation and Item-theta Correlation in Test Item Selection : A Simulation and Empirical Study. *Jurnal Penelitian dan Evaluasi Pendidikan*, *27*(2), 133–145. https://doi.org/https://doi.org/10.21831/pep.v27i2.61477

Mura, K. B., Sujito, Taufiq, A., Irawan, I. D. A., & Basri, N. A. (2025). Empirical Validation of a Conceptual Understanding Test on Dynamic Electricity for High School Students. *Jurnal Penelitian Pendidikan IPA*, *11*(11), 229–235. https://doi.org/10.29303/jppipa.v11i11.12428

Ramos, J. H., & Araya, R. (2025). Do School Activities Foster Creative Thinking ? An Analysis of PISA Results. *Education Sciences*, *15*(133), 1–27. https://doi.org/https://doi.org/10.3390/educsci15020133

Sabila, A. N., Wati, M., Dewantara, D., & Lasiani. (2024). Identification of Physics Learning Models in Improving Learners' Creative Thinking Skills: A Systematic Literature Review. *Jurnal pembelajaran fisika*, *12*(2), 112–128. https://doi.org/http://dx.doi.org/10.23960/jpf.v12.n2.202404

Sari, E. D. P., Trisnawati, R. K., Agustina, M. F., Adiarti, D., & Noorashid, N. (2023). Assessment of Students' Creative Thinking Skill on the Implementation of Project-Based Learning. *International Journal of Language Education*, *7*(3), 414–428.

https://doi.org/https://doi.org/10.26858/ijole.v7i3.38462

Sari, I. K., Handayani, R. D., & Supriadi, B. (2024). Students' Creative Thinking Skills on the Material of Motion Dynamics Through STEM-PjBL. *Jurnal Riset dan Kajian Pendidikan Fisika*, *11*(1), 41–48. https://doi.org/10.12928/jrkpf.v11i1.656

Setyaedhi, H. S. (2024). Comparative Test of Cronbach's Alpha Reliability Coefficient, Kr-20, Kr-21, And Split-Half Method. *Journal of Educational Research and Evaluation*, *8*(1), 47–57. https://doi.org/https://doi.org/10.23887/jere.v8i1.68164

Subhaktiyasa, P. G. (2024). Evaluasi Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif : Sebuah Studi Pustaka. *Journal of Education Research*, *5*(4), 5599–5609. https://doi.org/https://doi.org/10.37985/jer.v5i4.1747

Suherman, S., & Vidakovich, T. (2024). Role of Creative Self-efficacy and Perceived Creativity as Predictors of Mathematical Creative Thinking : Mediating Role of Computational Thinking. *Thinking Skills and Creativity journal*, *53*, 1–14. https://doi.org/10.1016/j.tsc.2024.101591

Suratno, Komaria, N., Yushardi, Dafik, & Wicaksono, I. (2019). The Effect of Using Synectics Model on Creative Thinking and Metacognition Skills of Junior High School Students. *International Journal of Instruction*, *12*(3), 133–150. https://doi.org/https://doi.org/10.29333/iji.2019.1239a

Susilowati, N. E., Muslim, Efendi, R., & Samsudin, A. (2022). Pisa 2021 Creative Thinking Instrument for Students : Physics Teachers' Perceptions. *Indonesian Journal of Science and Mathematics Education*, *05*(July), 194–209. https://doi.org/10.24042/ijsme.v5i1.12439

Tumangger, M. H., Cahyaningrum, R., Maghfiroh, A., & Sahrul. (2025). Validity and Reliability of Creative Thinking Skills Assessment Instruments in Mathematics Subjects for Elementary School. *Jurnal Pendidikan Matematika dan Matematika*, *8*(2), 448–460. https://doi.org/https://doi.org/10.30606/absis.v8i2.3316

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.

Widyastuti, N. L. M. Y., & Jayanta, I. N. L. (2024). Design Thinking-Based Learning Assessment Guidelines to Improve Creative Thinking Skills. *International Journal of Elementary Education*, *8*(3), 425–437. https://doi.org/https://doi.org/10.23887/ijee.v8i3.74596

Zakariya, Y. F. (2022). Cronbach's Alpha in Mathematics Education Research: Its Appropriateness , Overuse , and Alternatives in Estimating Scale Reliability. *Frontiers in Psychology*, 1–6. https://doi.org/10.3389/fpsyg.2022.1074430