Journal of Advanced Sciences and Mathematics Education
Volume 5, Issue 1, 97 - 112.
e_ISSN: 2798-2351
DOI: 10.58524/jasme.v5i1.558

Creative problem-solving assessment in fluid mechanics for senior high school students: Instrument validation and reliability analysis using the rasch model

Rahma Diani*
Universitas Lampung
INDONESIA

Viyanti
Universitas Lampung,
INDONESIA

Tri Jalmo
Universitas Lampung,
INDONESIA

Dewi Lengkana
Universitas Lampung,
INDONESIA

Intan Erliana
UIN Raden Intan Lampung,
INDONESIA

Article Info

Article history:

Received: January 20, 2025
Revised: May 06, 2025
Accepted: June 12, 2025

Keywords:

Assessment Instrument;
Creative Problem-Solving;
Physics Education;
Rasch Model;
Validity and Reliability.

Abstract

Background: Creative Problem-Solving (CPS) is a critical skill in physics education, enabling students to address real-world challenges through innovative and practical approaches. Despite its importance, there is a limited number of reliable and valid instruments specifically designed to assess CPS skills within the context of physics education.
Aim: This study aims to develop and prove the validation of a CPS assessment instrument using the Rasch model to ensure its reliability and accuracy in measuring CPS skills in physics education.
Method: The research followed the ADDIE development model, which includes analysis, design, development, implementation, and evaluation phases. The instrument, consisting of 20 essay items aligned with CPS indicators, was validated by three experts in physics education. It was then tested on 35 Grade XI students in Lampung Province. Data analysis was performed using the Rasch model through Winsteps software, focusing on fit analysis, reliability, item difficulty distribution, and dimensionality assessment.
Results: The instrument demonstrated strong internal consistency, confirming its reliability for assessing CPS skills in physics education. The construct validity was supported by fit analysis, indicating that most items functioned appropriately. However, some items exhibited potential bias and required revision. The dimensionality analysis confirmed that the instrument effectively measured a single underlying construct, ensuring its psychometric robustness. These findings suggest that the developed CPS assessment instrument is a reliable and valid tool for evaluating students' creative problem-solving abilities in high school physics, particularly in fluid mechanics.
Conclusion: This study successfully developed and validated a CPS assessment instrument for high school physics education, demonstrating strong reliability and construct validity. The instrument effectively measures CPS skills, though some items require refinement to ensure fairness and accuracy. Future research should focus on further improving item quality and testing the instrument in diverse educational settings to enhance its applicability and generalizability.

To cite this article: Diani, R., Viyanti., Jalmo, T., Lengkana, D., & Erliana, I. (2025). Creative Problem-Solving Assessment in Fluid Mechanics for Senior High School Students: Instrument Validation and Reliability Analysis Using the Rasch Model. *Journal of Advanced Sciences and Mathematics Education, 5*(1), 97 - 112.

INTRODUCTION

Creative problem-solving (CPS) skills play a pivotal role in physics education, particularly in this era of rapid advancements in science and technology. CPS serves as a fundamental competency to prepare students for the complexities of an ever-evolving real world, where they are expected to identify problems, analyze them from diverse perspectives, and formulate innovative yet practical

* **Corresponding author:**
Rahma Diani, UIN Raden Intan Lampung, INDONESIA
rahmadiani@radenintan.ac.id

solutions (Aytekin & Topçu, 2024; Puccio et al., 2022). Within the framework of scientific literacy, CPS in physics learning encourages students not only to gain a deep understanding of physical concepts but also to connect these concepts to real-world phenomena through approaches that involve analysis, synthesis, and evaluation (Bao & Koenig, 2019; DeHaan, 2009; Shakhman & Barak, 2019). Additionally, CPS has been shown to enhance critical and reflective thinking, which are essential for the development of scientific competence (Fathonah et al., 2023; García-Carmona, 2023).

Numerous studies have highlighted the significance of CPS in fostering students' higher-order thinking, particularly in science and engineering education (Shin et al., 2025; Song et al., 2025; Xia et al., 2025). CPS-based assessments encourage students to engage with complex tasks that demand creativity, reasoning, and adaptability (Farida et al., 2024a; Pears et al., 2025). Furthermore, the use of the Rasch model has become increasingly prominent in the validation of instruments designed to measure creative and critical thinking in physics (Kassiavera et al., 2024; Naingalis et al., 2023). Several researchers have explored CPS-oriented assessments in topics such as electricity, optics, and quantum mechanics (Bitzenbauer et al., 2022; Testa et al., 2020; Widyaningsih et al., 2021), while Rasch-based refinements have been utilized to improve measurement quality in STEM education (Purnami et al., 2023; Ringo et al., 2021). Despite these promising efforts, very few studies have specifically addressed CPS assessment in senior high school physics, particularly in fluid mechanics—a topic known for its abstract and conceptual complexity (Marfu'i et al., 2019; Yusuf et al., 2020). Moreover, existing instruments are often either too general for science education or developed for higher education contexts, leading to a significant gap in tools aligned with the high school physics curriculum. Although ICT-based assessments have shown potential in evaluating CPS (Farida et al., 2024b), their psychometric robustness in specific domains such as fluid mechanics remains underexplored.

Given this gap, the development of a robust CPS assessment instrument in physics education is crucial to ensure reliable and valid measurement results. Such instruments are essential for accurately identifying students' CPS levels, enabling educators to design more effective strategies for enhancing these skills (Treffinger, 1995; Ying & Tiemann, 2024). Assessing CPS in physics also necessitates integrating the relationship between physics concepts and realistic problem-solving contexts to ensure that the assessment outcomes reflect abilities relevant to students' future needs (Burkholder et al., 2020; Maries & Singh, 2023).

To evaluate the developed CPS instrument, the Rasch model has been widely adopted due to its capacity to provide in-depth analysis regarding the validity and reliability of assessment tools (Baghaei, 2008; Boone & Staver, 2020). The Rasch model allows for a comprehensive evaluation of both student abilities and item difficulty levels within the instrument, thereby offering robust evidence for construct validity and measurement reliability (Medvedev & Krägeloh, 2022; Testa et al., 2020). This model can map how effectively the instrument's items measure students' CPS skills in physics while also identifying items that may require revision (Planinic et al., 2019). Consequently, employing the Rasch model in this study is vital to ensure that the developed CPS instrument consistently yields accurate and reliable results.

This study aims to develop and validate a CPS assessment instrument specifically designed for senior high school physics education, with a particular focus on fluid mechanics. By employing the Rasch model, this research provides a comprehensive perspective on the instrument's effectiveness in measuring students' CPS skills, as well as evaluating its consistency and accuracy. Therefore, this study contributes significantly to the field of educational assessment, particularly in the development of psychometrically sound instruments tailored to physics learning in secondary education.

METHOD

The instrument development process followed the ADDIE model (Branch, 2009; Diani & Hartati, 2018), as illustrated in Figure 1. The analysis phase ensured that the instrument was aligned with real-world needs and conditions. During the design phase, 20 essay items were developed to measure students' CPS skills in the context of fluid mechanics. The instrument incorporated CPS indicators, including objective finding, fact finding, problem finding, idea finding, solution finding, and acceptance finding (Chen et al., 2021; Diani et al., 2019; Fiteriani et al., 2021; Mitchell & Kowalik, 1979). Subsequently, the items were validated by three experts in physics education who assessed their appropriateness, clarity, and relevance.

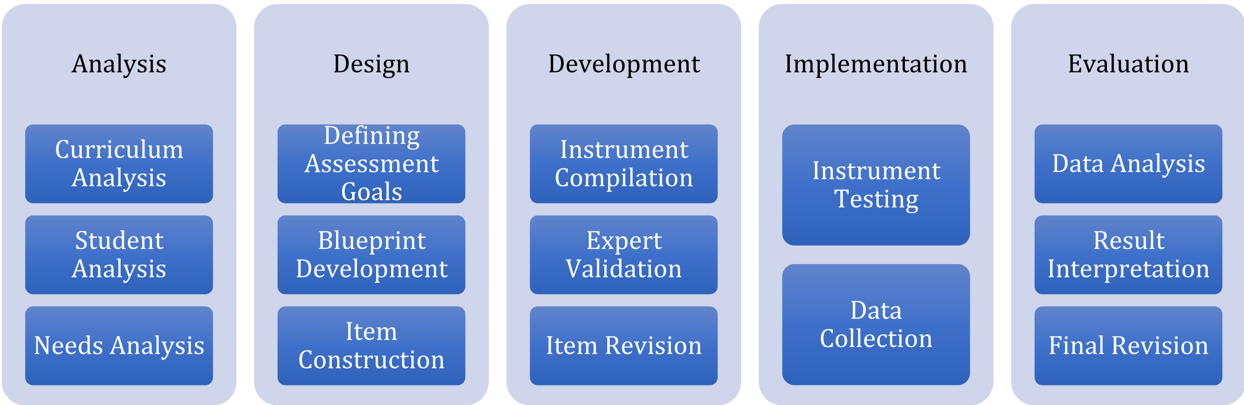

Figure 1. Research phases

After expert validation, the instrument was tested on 35 Grade XI students from a senior high school in Lampung Province who had prior knowledge of fluid mechanics. The validation process involved three expert validators, consisting of two physics education specialists and one assessment and evaluation expert. These validators were selected based on their academic qualifications, extensive experience in educational assessment, and their research background in problem-solving and physics education. Participants for the trial were selected through purposive sampling to ensure their familiarity with the tested material. The students answered the 20 essay items within a specified timeframe. Data collected from this trial were used for further analysis. The study adhered to ethical considerations, including obtaining school approval and participant consent. Confidentiality was strictly maintained, and participation was voluntary.

The trial data were analyzed using the Rasch model with the aid of Winsteps software (Bond, 2015; Linacre, 2002). The analysis included the following components:

- Verification of the Unidimensionality Assumption: This step ensures that the instrument measures only a single primary construct, namely Creative Problem-Solving (CPS), in accordance with the unidimensionality principle of the Rasch model (Ambrosio et al., 2020; Huang et al., 2023)
- Fit Analysis: This test evaluates the conformity of test items to the Rasch model by examining the values of Outfit MNSQ, Outfit ZSTD, and Pt-Measure Corr (Christensen et al., 2019; Elder, 2024). An item is considered misfitting if it does not meet the following criteria: Outfit mean-square residual (MNSQ): $0.5 < y < 1.5$, Outfit standardized mean-square residual (ZSTD): $-2.0 < Z < +2.0$, and Point Measure Correlation (Pt-Measure Corr): $0.4 < x < 0.8$.
- Rating Scale: This analysis examines the effectiveness of the response categories used in the CPS assessment. It evaluates whether rating scales function as intended by analyzing category thresholds, response distribution, and step calibration to ensure a meaningful progression in scoring.

- Person-Item Map (Wright Map) Analysis: This analysis examines the alignment between item difficulty levels and students' CPS abilities to ensure the instrument effectively measures the intended construct.
- Item Separation Analysis: This test assesses how well the instrument differentiates items into different difficulty levels (e.g., easy, moderate, difficult) within the CPS framework.
- Person Separation Analysis: This test evaluates how effectively the instrument categorizes respondents based on their CPS ability levels (e.g., low, moderate, high).
- Reliability Analysis: Person and item reliability indices were calculated to evaluate the instrument's consistency. High reliability indicates that the instrument provides consistent measurements.
- Item Difficulty Analysis: The difficulty level of each item was determined to ensure a balanced distribution aligned with students' abilities.
- Differential Item Functioning (DIF) Analysis: This analysis detects potential bias in specific items against particular groups (gender-based bias). Ensuring fairness in the instrument is crucial for maintaining validity across diverse respondent groups.

Based on the analysis results, items that did not fit the Rasch model or exhibited poor fit indices were revised or removed to enhance the instrument's validity and reliability. The rating scale analysis ensured that response categories functioned effectively, contributing to meaningful score interpretation. The person-item map analysis confirmed the alignment between item difficulty levels and respondents' CPS abilities, while item and person separation indices demonstrated the instrument's capability to differentiate both item difficulty and student ability levels. Additionally, reliability analysis confirmed the consistency of the instrument, item difficulty analysis ensured a balanced distribution of cognitive challenges, and DIF analysis verified the absence of significant bias across respondent groups. Taken together, these analyses validate the CPS assessment instrument as a robust and equitable tool for measuring creative problem-solving skills in physics education.

RESULTS AND DISCUSSION

Results

Curriculum analysis underscores the importance of developing higher-order thinking skills, including Creative Problem-Solving (CPS), in physics education. This process is conducted to ensure that the material aligns with learning objectives and meets students' needs. A thorough curriculum analysis includes examining learning outcomes, analyzing vertical and horizontal competency relationships, and evaluating material relevance to students' characteristics and local contexts. Ensuring this alignment is essential for designing assessments that accurately measure CPS skills. However, the implementation of CPS-focused learning in real-world classrooms remains suboptimal, necessitating well-designed assessment instruments to evaluate students' CPS development effectively. Observations and interviews with students revealed that many struggle to apply physics concepts when solving problems requiring creativity. This difficulty is attributed to a lack of practice and assessments that emphasize CPS development. Students are often accustomed to routine questions and are rarely exposed to challenges that encourage creative thinking. By ensuring that curriculum materials and assessments align with both cognitive development and contextual learning, educators can enhance students' ability to engage in creative problem-solving in physics.

Interviews with physics educators further highlighted gaps in the current CPS assessment practices. Teachers expressed the need for valid and reliable instruments to effectively evaluate students' CPS abilities, which could assist in designing better teaching strategies and providing constructive feedback. Based on the needs analysis and curriculum analysis results, a 20-item question blueprint was developed, aligning with key CPS indicators and ensuring that the questions addressed both conceptual understanding and real-world problem-solving skills. The decision to use

20 items was based on the balance between comprehensive assessment coverage and practical feasibility for classroom implementation.

The instrument underwent expert validation involving three specialists in physics education and educational assessment. Validators provided comments and suggestions regarding question clarity, alignment with CPS indicators, and appropriateness for students' cognitive levels. Some items required revisions before finalization, particularly in terms of wording and contextual relevance. Following expert feedback, modifications were made to improve the clarity, relevance, and appropriateness of specific items. Additionally, clear scoring guidelines were refined to enhance consistency in assessment. As a result, the validated instrument serves as a comprehensive and accurate tool for assessing CPS skills in physics education.

Following revisions, the instrument was tested on a representative sample of students. Before conducting further analysis using the Rasch model, it is essential to ensure that the instrument meets the assumptions of unidimensionality and local independence. The unidimensionality test was conducted using Winsteps, focusing on the dimensionality output to assess the proportion of variance explained by the model. The results are presented in Figure 2.

```
TABLE 23.0 INPUR RASCH BUNDA RAHMA.xlsx          ZOU968WS.TXT  Oct 24 2024 15:29
INPUT: 35 PERSON  20 ITEM  REPORTED: 35 PERSON  20 ITEM  6 CATS WINSTEPS 5.7.3.0
--------------------------------------------------------------------------------

     Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = ITEM information units
                                           Eigenvalue   Observed   Expected
Total raw variance in observations     =      27.1856 100.0%         100.0%
  Raw variance explained by measures   =       7.1856  26.4%          26.8%
    Raw variance explained by persons  =       5.2172  19.2%          19.4%
    Raw Variance explained by items    =       1.9685   7.2%           7.3%
  Raw unexplained variance (total)     =      20.0000  73.6% 100.0%   73.2%
    Unexplned variance in 1st contrast =       3.0909  11.4%  15.5%
    Unexplned variance in 2nd contrast =       2.6543   9.8%  13.3%
    Unexplned variance in 3rd contrast =       2.1398   7.9%  10.7%
    Unexplned variance in 4th contrast =       2.0122   7.4%  10.1%
    Unexplned variance in 5th contrast =       1.6871   6.2%   8.4%
```

Figure 2. Unidimensionality analysis of the CPS assessment instrument

The analysis indicates that the Rasch model explains 26.4% of the total variance in the CPS assessment instrument, exceeding the 20% threshold required to meet the unidimensionality assumption (Islam et al., 2020). Given the polytomous scoring system, this value is sufficient to confirm that the instrument primarily measures creative problem-solving skills in physics education. The first contrast in unexplained variance is 11.4%, which, while indicating some residual variance, remains below the 15% threshold, suggesting that potential secondary dimensions do not significantly affect the instrument's measurement focus (Li et al., 2024; Park, 2021). These findings confirm that the CPS assessment instrument meets the unidimensionality assumption, validating its use in physics education assessments.

The item fit analysis demonstrates that the majority of test items align well with the Rasch model expectations. Most items exhibit appropriate fit, indicating that they function effectively in measuring the intended construct. This suggests that the test as a whole has strong psychometric properties, with well-calibrated items across varying difficulty levels.

Table 1. Item fit Analysis

Items	MNSQ	ZSTD	PTMA	MNSQ	ZSTD	PTMA	Conclusion
P1	1.4347	1.8414	0.3504	fit	fit	misfit	FIT
P2	0.2709	-4.8397	0.6761	misfit	misfit	fit	FIT
P3	1.4786	1.9715	0.4828	fit	fit	fit	FIT
P4	1.1679	0.7912	0.3929	fit	fit	misfit	FIT
P5	0.79	-0.9892	0.4291	fit	fit	fit	FIT
P6	0.7065	-1.4493	0.7079	fit	fit	fit	FIT

P7	0.7957	-0.9892	0.561	fit	fit	fit	FIT
P8	0.5245	-2.7395	0.557	fit	misfit	fit	FIT
P9	1.1057	0.5411	0.6431	fit	fit	fit	FIT
P10	0.8365	-0.7592	0.3699	fit	fit	misfit	FIT
P11	1.1576	0.7912	0.5401	fit	fit	fit	FIT
P12	1.1932	0.9512	0.387	fit	fit	misfit	FIT
P13	1.0147	0.141	0.5132	fit	fit	fit	FIT
P14	1.0645	0.3711	0.4396	fit	fit	fit	FIT
P15	1.4764	2.0515	0.0848	fit	misfit	misfit	FIT
P16	0.8905	-0.4691	0.4342	fit	fit	fit	FIT
P17	0.7957	-0.9892	0.5183	fit	fit	fit	FIT
P18	1.2727	1.2813	0.3584	fit	fit	misfit	FIT
P19	1.0006	0.071	0.4561	fit	fit	fit	FIT
P20	1.1887	0.9312	0.4869	fit	fit	fit	FIT

Two items, P2 and P15, stand out as they only meet one of the established fit criteria. While they are still classified as fit, their marginal alignment with the model warrants further investigation. These items may require closer scrutiny to determine if their response patterns indicate inconsistencies in how different groups of participants interact with them. Such inconsistencies could arise from contextual bias or differences in interpretation across participant demographics. To ensure the fairness and validity of the assessment, a follow-up Differential Item Functioning (DIF) analysis is recommended. This analysis will help identify whether certain items exhibit bias toward specific participant subgroups, such as gender. If a significant DIF effect is detected, it may be necessary to revise or replace the affected items to maintain the overall test integrity.

Despite these minor concerns, the overall distribution of item fit results suggests that the instrument is well-structured. Most items effectively differentiate participants based on ability levels, supporting the test's reliability and measurement precision. The results provide a strong foundation for further refinement, ensuring that the assessment remains a robust tool for evaluating participants' skills in a fair and valid manner.

The evaluation of the rating scale in the scoring rubric was conducted by examining the Andrich Thresholds to determine whether the score categories effectively differentiate participants' ability levels. Ideally, the threshold distance should range between 1.4 and 5.0 logit. The analysis results indicate that some categories have excessively small threshold distances, particularly in the transitions from rating 2 to 3 (0.80 logit), 3 to 4 (0.55 logit), and 4 to 5 (0.13 logit), necessitating their combination. Meanwhile, rating categories 0, 1, and 2 meet the criteria and can be retained.

Table 2. The evaluation of the rating scale in the scoring rubric

Rating Transition	Lower Threshold	Upper Threshold	Andrich Threshold Distance	Recomendation
Rating 0 to 1	0	-3,43	3,43	Retain
Rating 1 to 2	-3,43	-0,05	-3,38	Retain
Rating 2 to 3	-0,05	0,75	0,80	Merge
Rating 3 to 4	0,75	1,3	0,55	Merge
Rating 4 to 5	1,3	1,43	0,13	Merge

The category probability curve further supports this recommendation, as only categories 0, 1, 2, and 5 exhibit distinct distribution peaks (see figure 3), while categories 3 and 4 do not form sufficiently clear probability patterns. Therefore, it is recommended to modify the rating scale from six categories to four, merging categories 3, 4, and 5 into a single category (rating 3). This adjustment is expected to enhance the validity of the rating scale in accurately mapping students' abilities.

```
         CATEGORY PROBABILITIES: MODES - Andrich thresholds at intersections
P      -+------+------+------+------+------+------+------+------+-
R  1.0 +                                                         +
O      |                                                         |
B      |                                                         |
A      |0                                                        |
B   .8 + 00                                                     5+
I      |   00                                                 55 |
L      |     0           111111111                           5   |
I      |      00        11         11                       5    |
T   .6 +        0     11             11                    5     +
Y      |         00 11                 1                  5      |
    .5 +           *                    1                5       +
O      |         11 0                    1              5        |
F   .4 +        1    00                   *222         5         +
       |      11       0                 22 1  22     5          |
R      |     1          00             22    1  3**333*44444     |
E      |   11             0          22       *3   *4*33    44   |
S   .2 + 11                00     222       33 11 4 *   3     444 +
P      |1                    00*2        33    4* 5 22  33       4|
O      |                    222 000    33    44 5*1   22  333    |
N      |             2222222      3****0 4444555   111  222  3333|
S   .0 +************************************000000000***********+
E      -+------+------+------+------+------+------+------+------+-
       -5     -4     -3     -2     -1      0      1      2      3
        PERSON [MINUS] ITEM MEASURE
```

Figure 3. Graph of Andrich Thresholds at category intersections

The Wright Map provides a comprehensive overview of the balance between respondents' ability levels and item difficulty within the same scale. The respondent distribution on the left side indicates that most participants have a relatively good ability level, with logits ranging between 0 and 2. Meanwhile, on the right side, the item distribution reflects a sufficiently broad range of difficulty levels. Most items are well-distributed, indicating that the instrument effectively measures respondents' abilities.

```
MEASURE                 PERSON - MAP - ITEM
                            <more>|<rare>
    3                             +
                                  |
                                  |
                                  |
                                  |
                             30P  |
                                  |b
                                  |
    2                             +
                                 T|
                             05P  |
                             03L  |
                        06L  21L  |
                                 S|
      01P  02L  04P  10P  20L     |
                07P  16P  17L     |
    1           09P  18P  19P  +
                             15P M|
                08P  11L  12P  |  P13
           14P  24P  27L  33P  |
      23P  25P  29L  32P  34L  |T
      13L  22L  26P  31P  35L S|  P4
                          28P  |S P2      P3      P6      P9
                               |  P15
    0                         T+M P10     P11     P17     P7
                               |  P12     P14     P18     P19     P8
                               |S P20
                               |  P16     P5
                               |T P1
                               |
                               |
                               |
   -1                          +
                          <less>|<freq>
```

Figure 4. Wirght Map (Person-Item map)

At the M (+2 SD) level, there is one participant (30P), indicating that this individual possesses exceptionally high ability compared to the others. No items are present at this level, suggesting that there are no questions sufficiently difficult to differentiate individuals with such extreme abilities. In the range of S (+1 SD) to M (Mean), several participants (05P, 03L, 06L, 21L) exhibit high ability levels. Some items are also positioned in this range, indicating that these questions are relatively difficult and can only be answered correctly by participants with higher abilities. Between M (Mean) and -S (-1 SD), participants and items are more evenly distributed. This suggests that most participants fall within this ability range, and the majority of items also exhibit a moderate level of difficulty. At ≤ -S (-1 SD), some participants demonstrate lower abilities compared to others. Several items are also located within this range, indicating that these questions are relatively easier and can be answered correctly by most participants.

Overall, the Wright Map illustrates a relatively balanced distribution of participant abilities, with items spread across different difficulty levels. However, there is a slight imbalance at the highest ability level, where a participant with exceptionally high ability lacks sufficiently challenging items to assess their full potential.

Table 3. Person and Item Reliability and Separation Statistics

	Rerata Logit (SD)	Separation	Reliability	Alpha Cronbach
Person	0,89 (0,48)	1,92	0,79	0,81
Item	0,00 (0,28)	1,35	0,65	

The analysis of person and item separation provides insight into the effectiveness of the instrument in distinguishing between different levels of ability. The person separation index of 1.92 suggests that the instrument is capable of categorizing participants into approximately two distinct ability groups. This indicates a reasonable level of differentiation, meaning that the test can effectively distinguish between higher- and lower-ability participants. Additionally, the person reliability of 0.79 supports the consistency of this categorization, reinforcing the instrument's stability in measuring participants' abilities. The Cronbach's Alpha value of 0.81 further confirms that the test demonstrates good internal consistency, indicating that the items work well together in assessing the intended construct.

On the item side, the item separation index of 1.35 suggests that the test items vary in difficulty, but the distinction is not as strong as that observed in person separation. This indicates that while there is some spread in item difficulty, additional refinement may enhance the ability of the test to differentiate more effectively across a wider range of ability levels. The item reliability of 0.65 suggests a moderate level of consistency in item difficulty, meaning that while the items are generally stable, further calibration may be beneficial to strengthen the test's overall structure.

Overall, these reliability indicators suggest that the instrument is well-developed and provides meaningful differentiation among participants. The KR-10 reliability analysis, reflected in the Cronbach's Alpha value, confirms that the test is internally consistent and appropriately structured. While the instrument already demonstrates strong reliability, slight refinements—particularly in item difficulty calibration—could further enhance its effectiveness in differentiating participants across ability levels. These findings provide a positive outlook on the test's capability to measure competencies accurately while also highlighting opportunities for refinement.

The distribution of item difficulty in this instrument demonstrates a well-balanced range of difficulty levels, with items spanning from "Very Easy" to "Very Difficult." This spread allows the instrument to effectively measure participants' abilities across various levels. Several items fall into the "Very Easy" category, including P1 (-0.47), P5 (-0.43), and P16 (-0.38). These items are likely to be answered correctly by most participants, indicating their low difficulty level. The presence of such items is crucial to ensure that lower-ability participants can still engage with the test effectively.

Table 4. The distribution of item difficulty

Categroy	Range	Items
Very Difficult	> 1 SD	P3, P4, P13,
Difficult	Mean s/d 1 SD	P2, P6, P9, P10, P15
Easy	-1 SD s/d mean	P7, P8, P11, P12, P14, P17, P18, P19, P20
Very Easy	< -1 SD	P1, P5, P16

Conversely, some items are categorized as "Very Difficult," such as P3 (0.3), P4 (0.33), and P13 (0.7). These items are designed to challenge high-ability participants and differentiate them from those with moderate or lower abilities. However, an excessive number of high-difficulty items may lead to an imbalance that could hinder the assessment of lower-ability participants. The "Difficult" category includes items such as P2 (0.2), P6 (0.2), P9 (0.28), P10 (0.06), and P15 (0.09). These items serve as intermediaries between easier and harder items, contributing to a more even distribution of difficulty levels within the instrument.

Most items are classified as "Easy," including P7 (-0.03), P8 (-0.12), P11 (-0.05), P12 (-0.1), P14 (-0.15), P17 (-0.03), P18 (-0.08), P19 (-0.08), and P20 (-0.24). These items help ensure that participants with average ability levels can still demonstrate their understanding without facing excessive challenges. Overall, the distribution of item difficulty exhibits a good balance between easy and difficult items. However, further analysis is necessary to ensure that the "Very Difficult" and "Very Easy" items do not overly dominate, thus maintaining the instrument's ability to assess participants' abilities optimally.

The results of the DIF analysis indicate that most items exhibit no significant bias between groups. This conclusion is based on the probability (p-value) exceeding the 0.05 threshold for almost all items. However, item P13 has a p-value of 0.0357, which is below 0.05, suggesting potential bias. Additionally, item P15 has a p-value of 0.0511, which is close to the threshold, warranting further attention.

PERSON CLASSES	SUMMARY DIF CHI-SQUARED	D.F.	PROB.	BETWEEN-CLASS/GROUP UNWTD MNSQ	ZSTD	ITEM Number	Name
2	.7432	1	.3886	.7841	.31	1	P1
2	.5704	1	.4501	.5961	.14	2	P2
2	1.5011	1	.2205	1.6063	.83	3	P3
2	.0000	1	1.0000	.0085	-1.22	4	P4
2	.2307	1	.6310	.2401	-.33	5	P5
2	.0099	1	.9209	.0172	-1.10	6	P6
2	.0474	1	.8276	.0485	-.88	7	P7
2	.2985	1	.5848	.3116	-.21	8	P8
2	.0290	1	.8648	.0296	-.99	9	P9
2	.0000	1	1.0000	.0032	-1.34	10	P10
2	.1697	1	.6804	.1755	-.46	11	P11
2	1.7786	1	.1823	1.9232	.99	12	P12
2	4.4091	1	.0357	5.0706	1.99	13	P13
2	.3332	1	.5638	.3427	-.17	14	P14
2	3.8055	1	.0511	4.3086	1.80	15	P15
2	.3171	1	.5734	.3292	-.19	16	P16
2	1.2727	1	.2593	1.3587	.70	17	P17
2	.1027	1	.7486	.1068	-.64	18	P18
2	1.5049	1	.2199	1.6161	.84	19	P19
2	1.2088	1	.2716	1.2963	.66	20	P20

Figure 5. Differential Item Functioning (DIF) analysis results

The DIF measure plot further supports these findings, showing a noticeable deviation for items P13 and P15 compared to others. The graph illustrates significant differences in responses between male and female participants, reinforcing the indication of potential bias in these items. These results suggest that item P13 demonstrates a statistically significant DIF, while item P15 is on the borderline.

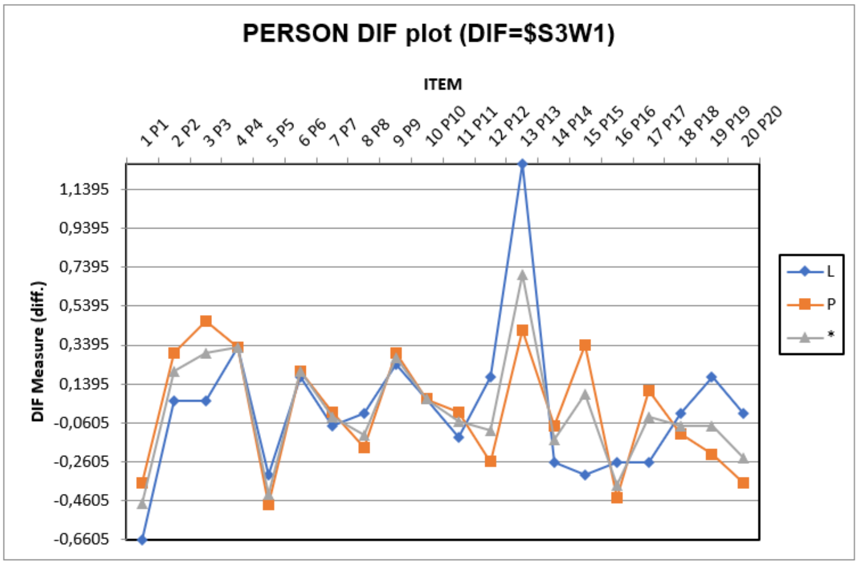

Figure 6. Differential Item Functioning (DIF) analysis graph by gender (L: Male, P: Female).

Following this analysis, a revision of item P13 is necessary to eliminate any potential bias. Item P15 also requires further review to ensure that it functions fairly across different respondent groups. Additional investigations can be conducted to explore the source of bias in these items, such as qualitative analysis or expert review to refine item wording and ensure content neutrality.

Discussion

The results of this study indicate that the developed creative problem-solving (CPS) assessment instrument possesses strong psychometric properties, particularly in terms of validity and reliability. The Cronbach's Alpha coefficient of 0.81 reflects high internal consistency, exceeding the recommended threshold for educational instruments (Barbera et al., 2021; Taber, 2018). Additionally, the item fit statistics confirmed that the items align well with the Rasch model, providing evidence of strong construct validity. These results suggest that the instrument is capable of measuring students' CPS abilities in fluid mechanics reliably and meaningfully.

Despite the instrument's overall strength, several areas were identified for refinement. The analysis of rating scale functionality revealed that some scale categories were underutilized, which may suggest ambiguity or limited relevance to respondents. As recommended by Adams and Wieman (2011) and Kalkbrenner (2021), optimizing these categories is essential to enhance response sensitivity and ensure a more nuanced assessment of CPS levels. Furthermore, Differential Item Functioning (DIF) analysis identified certain items—particularly item P13—as exhibiting potential bias across different student subgroups. Such bias poses a risk to measurement fairness and may compromise the generalizability of findings. Addressing item bias through item revision or recalibration is therefore necessary to uphold equity and validity in assessment (Robertson & Larki, 2019).

Another critical finding concerns the distribution of item difficulty. Although the instrument covers a range of student abilities, a number of items were found to be too easy and lacked sufficient discriminatory power, especially for higher-achieving students. This can limit the diagnostic value of the instrument in differentiating between moderate and advanced CPS abilities. Item-level analyses—such as difficulty and discrimination indices—should guide future revisions to ensure that the assessment captures the full spectrum of student proficiency (Maba et al., 2017; Smith & McCarthy, 1995).

The present findings are consistent with those of previous studies investigating CPS assessment in science education. Ulya et al. (2024) reported a reliability coefficient of 0.82 for a CPS instrument in middle school mathematics learning, while Viyanti et al. (2022) obtained a coefficient of 0.80 in a project-based physics learning context. Similarly, Agnezi and Festiyed (2023) demonstrated empirical validity and reliability for instruments assessing problem-solving skills in fluid mechanics and thermodynamics. These comparable results support the assertion that well-constructed CPS instruments can effectively measure complex cognitive abilities in science learning.

Taken together, these findings highlight the utility of the developed CPS instrument for assessing students' creative problem-solving skills in high school physics, specifically within the domain of fluid mechanics. They also underscore the importance of continuous refinement—through scale optimization, bias detection, and difficulty balancing—to improve the instrument's psychometric robustness. This study contributes not only to the practical application of CPS assessment in physics education but also provides a methodological foundation for the development of valid and reliable instruments using Rasch modeling tailored to domain-specific cognitive demands.

Implications

The findings of this study have significant implications for both educational theory and practice, particularly in the development of CPS assessment instruments for physics education. The high reliability and satisfactory validity of the instrument reinforce the theoretical framework that CPS skills can be measured objectively and consistently, aligning with prior research. From a practical perspective, this instrument provides educators with a structured tool to assess students' creative problem-solving abilities, allowing for more targeted instructional strategies. The assessment results can help teachers identify students' strengths and areas for improvement in CPS, enabling them to design interventions that foster higher-order thinking skills. Additionally, students' responses to the instrument provide insights into how they engage with complex, real-world physics problems, shedding light on their reasoning processes and problem-solving approaches.

Beyond classroom application, these findings serve as a valuable reference for policymakers in designing curricula that integrate CPS development as a core learning objective. By emphasizing CPS in physics education, schools can cultivate students' ability to approach scientific challenges with creativity and critical thinking. Furthermore, the validated instrument can be used for broader educational assessments, contributing to the refinement of national and international evaluation standards. Thus, this study contributes to the advancement of educational assessment methodologies and provides a solid empirical foundation for improving physics education by fostering creative problem-solving skills in students.

Limitations

While this study successfully developed a CPS assessment instrument with adequate reliability and validity, some limitations warrant consideration. First, the research sample was limited to one school, requiring caution when generalizing the findings to a broader population. Second, the Differential Item Functioning (DIF) analysis identified potential bias in certain items, such as P13, necessitating further evaluation to ensure fairness across diverse student groups. Finally, time and resource constraints restricted the study to a single round of testing. Repeated trials with more diverse samples would offer a more comprehensive understanding of the instrument's consistency and reliability. Despite these limitations, the findings significantly contribute to the development of CPS measurement tools in physics education.

Suggestions

Based on the findings of this study, it is recommended that the developed CPS assessment instrument be tested further on a broader and more diverse population, including students from different schools, regions, and educational backgrounds, to improve external validity and ensure generalizability. Additionally, items with potential bias, such as P13, should be revised and revalidated through expert reviews and cognitive interviews with students to ensure fairness and accuracy in measuring CPS across different demographic groups.

Future research should also explore the application of this instrument in various instructional settings, such as inquiry-based or project-based learning environments, to assess its adaptability and effectiveness in different teaching methodologies. Furthermore, longitudinal studies tracking students' CPS development over time using this instrument could provide deeper insights into its long-term reliability and its role in fostering creative problem-solving skills in physics education.

CONCLUSION

This study successfully developed a reliable and valid instrument for assessing CPS skills in physics education, aligning with the initial objectives outlined in the Introduction. The instrument's high reliability and satisfactory validity confirm its effectiveness in measuring CPS abilities among students. These findings contribute to both theoretical understanding and practical application in CPS assessment, providing educators with a structured tool to evaluate and enhance students' creative problem-solving skills in physics. To maximize its usability, educators are encouraged to integrate this instrument into classroom assessments and instructional strategies, particularly in inquiry-based or project-based learning settings. The instrument can also serve as a diagnostic tool to identify students' strengths and areas for improvement in CPS, enabling targeted interventions to support their problem-solving development.

Future research should focus on refining the instrument by addressing identified biases through iterative validation processes, including expert reviews and cognitive interviews with students. Additionally, expanding its implementation across different educational contexts, such as various grade levels and learning environments, will help establish its generalizability. Longitudinal studies examining the impact of CPS assessment on students' learning outcomes and problem-solving growth over time would further strengthen its effectiveness and applicability in science education. By continuously improving and adapting this instrument, it can become a valuable asset in fostering creative problem-solving skills in physics education, ultimately supporting students in developing the competencies required for scientific inquiry and real-world problem-solving.

AUTHOR CONTRIBUTIONS STATEMENT

RD developed the research concept, designed the assessment instrument, and drafted the initial manuscript. V supervised the overall research process, ensured academic integrity, and approved the final manuscript for publication. TJ contributed to developing the research methodology and provided critical feedback on manuscript revisions. DL supplied research resources and necessary equipment, as well as supported data collection. IE conducted data collection, analyzed data using the Rasch model, and contributed to interpreting the results.

REFERENCES

Adams, W. K., & and Wieman, C. E. (2011). Development and Validation of Instruments to Measure Learning of Expert-Like Thinking. *International Journal of Science Education*, *33*(9), 1289–1312. https://doi.org/10.1080/09500693.2010.512369

Agnezi, L. A. & Festiyed. (2023). Development of The Online Assessment Instrument for Fluid, Temperature, and Heat to measure the problem solving skills of high school students. *Journal of Physics: Conference Series*, *2582*(1), 012053. https://doi.org/10.1088/1742-6596/2582/1/012053

Ambrosio, L., Rodríguez-Blázquez, C., Ayala, A., & Forjaz, M. J. (2020). Rasch analysis of the living with chronic illness scale in Parkinson's disease. *BMC Neurology*, *20*(1). https://doi.org/10.1186/s12883-020-01926-7

Aytekin, A., & Topçu, M. S. (2024). Improving 6th Grade Students' Creative Problem Solving Skills Through Plugged and Unplugged Computational Thinking Approaches. *Journal of Science Education and Technology*, *33*(6), 867–891. https://doi.org/10.1007/s10956-024-10130-y

Baghaei, P. (2008). The Rasch model as a construct validation tool. *Rasch Measurement Transactions*, *22*(1), 1145–1146.

Bao, L., & Koenig, K. (2019). Physics education research for 21st century learning. *Disciplinary and Interdisciplinary Science Education Research*, *1*(1), 2. https://doi.org/10.1186/s43031-019-0007-8

Barbera, J., Naibert, N., Komperda, R., & Pentecost, T. C. (2021). Clarity on Cronbach's Alpha Use. *Journal of Chemical Education*, *98*(2), 257–258. https://doi.org/10.1021/acs.jchemed.0c00183

Bitzenbauer, P., Veith, J. M., Girnat, B., & Meyn, J.-P. (2022). Assessing Engineering Students' Conceptual Understanding of Introductory Quantum Optics. *Physics*, *4*(4), Article 4. https://doi.org/10.3390/physics4040077

Bond, T. (2015). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences, Third Edition* (3rd ed.). Routledge. https://doi.org/10.4324/9781315814698

Boone, W. J., & Staver, J. R. (2020). *Advances in Rasch Analyses in the Human Sciences*. Springer International Publishing. https://doi.org/10.1007/978-3-030-43420-5

Branch, R. M. (2009). *Instructional Design: The ADDIE Approach*. Springer US. https://doi.org/10.1007/978-0-387-09506-6

Burkholder, E. W., Miles, J. K., Layden, T. J., Wang, K. D., Fritz, A. V., & Wieman, C. E. (2020). Template for teaching and assessment of problem solving in introductory physics. *Physical Review Physics Education Research*, *16*(1), 010123. https://doi.org/10.1103/PhysRevPhysEducRes.16.010123

Chen, S. Y., Tsai, J. C., Liu, S. Y., & Chang, C. Y. (2021). The effect of a scientific board game on improving creative problem solving skills. *Thinking Skills and Creativity*, *41*(June), 100921. https://doi.org/10.1016/j.tsc.2021.100921

Christensen, K. S., Oernboel, E., Nielsen, M. G., & Bech, P. (2019). Diagnosing depression in primary care: A Rasch analysis of the major depression inventory. *Scandinavian Journal of Primary Health Care*, *37*(1). https://doi.org/10.1080/02813432.2019.1568703

DeHaan, R. L. (2009). Teaching Creativity and Inventive Problem Solving in Science. *CBE—Life Sciences Education*, *8*(3), 172–181. https://doi.org/10.1187/cbe.08-12-0081

Diani, R., & Hartati, N. S. (2018). Flipbook berbasis literasi Islam: Pengembangan media pembelajaran fisika dengan 3D pageflip professional. *Jurnal Inovasi Pendidikan IPA*, *4*(2), 234–244. https://doi.org/10.21831/jipi.v4i2.20819

Diani, R., Herliantari, Hesti, Irwandani, I., Saregar, A., & Umam, Rofiqul. (2019). *Search, Solve, Create, and Share (SSCS) Learning Model: The Impact on the Students Creative Problem-Solving Ability on the Concept of Substance Pressure*. *9*(1), 65–77. https://doi.org/10.26740/jpfa.v9n1.p65-77

Elder, H. (2024). Te Waka Kuaka, Rasch Analysis of a Cultural Assessment Tool in Traumatic Brain Injury in Māori. *Journal of Psychiatry and Psychiatric Disorders*, *08*(03). https://doi.org/10.26502/jppd.2572-519X0216

Farida, F., Alamsyah, Y. A., Anggoro, B. S., Andari, T., & Lusiana, R. (2024a). Rasch Measurement Validation of an Assessment Tool for Measuring Students' Creative Problem-Solving through the Use of ICT. *Pixel-Bit. Revista de Medios y Educación*, *71*, 83–106. https://doi.org/10.12795/pixelbit.107973

Farida, F., Alamsyah, Y. A., Anggoro, B. S., Andari, T., & Lusiana, R. (2024b). Rasch Measurement Validation of an Assessment Tool for Measuring Students' Creative Problem-Solving through the Use of ICT. *Pixel-Bit. Revista de Medios y Educación*, *71*, 83–106. https://doi.org/10.12795/pixelbit.107973

Fathonah, S., Cahyono, E., Iswari, R. S., Haryani, S., Sarwi, S., Lestari, N. H., & Kadarwati, S. (2023). Effects of Multirepresentation-Based Creative Problem-Solving Learning Model on Students' Critical Thinking and Diet Nutritional Quality. *Journal of Turkish Science Education*, *20*(4), 669–694. https://doi.org/10.36681/tused.2023.038

Fiteriani, I., Diani, R., Hamidah, A., & Anwar, C. (2021). Project-based learning through STEM approach: Is it effective to improve students' creative problem-solving ability and metacognitive skills in physics learning? *Journal of Physics: Conference Series*, *1796*(1), 012058. https://doi.org/10.1088/1742-6596/1796/1/012058

García-Carmona, A. (2023). Scientific Thinking and Critical Thinking in Science Education. *Science &Amp; Education*. https://doi.org/10.1007/s11191-023-00460-5

Huang, X., Ma, H., Wang, X., Zhong, J., Chen, Y., & Xu, M. (2023). Rasch analysis of the Hospital Anxiety and Depression Scale (HADS) in patients with major depression. *Anales De Psicología*, *39*(3). https://doi.org/10.6018/analesps.550551

Islam, A. Y. M. A., Gu, X., Crook, C., & Spector, J. M. (2020). Assessment of ICT in Tertiary Education Applying Structural Equation Modeling and Rasch Model. *Sage Open*, *10*(4). https://doi.org/10.1177/2158244020975409

Kalkbrenner, M. T. (2021). A Practical Guide to Instrument Development and Score Validation in the Social Sciences: The MEASURE Approach. *Practical Assessment, Research & Evaluation*, *26*. https://eric.ed.gov/?id=EJ1311131

Kassiavera, S., Suparmi, A., Cari, C., & Sukarmin, S. (2024). Application of Rasch Model in Two-Tier Test for Assessing Critical Thinking in Physics Education. *Journal of Baltic Science Education*, *23*(6), 1227–1242. https://doi.org/10.33225/jbse/24.23.1227

Li, X., Su, Z., Wang, L., Li, J., & Diao, Y. (2024). Psychometric evaluation of the pictorial scale of perceived movement skill competence in Chinese children: An item response theory analysis. *Child: Care, Health and Development*, *50*(4), e13275. https://doi.org/10.1111/cch.13275

Linacre, J. (2002). *A User''s Guide to Winsteps: Rasch-Model Computer Program*.

Maba, W., Perdata, I. B. K., & Astawa, I. N. (2017). Constructing Assessment Instrument Models for Teacher's Performance, Welfare and Education Quality. *International Journal of Social Sciences and Humanities*, *1*(3), 88–96. https://doi.org/10.29332/ijssh.v1n3.59

Marfu'i, L. N. R., Ilfiandra, & Nurhudaya. (2019). The analysis of critical thinking skills test in social-problems for physics education students with Rasch Model. *Journal of Physics: Conference Series*, *1280*(5), 052012. https://doi.org/10.1088/1742-6596/1280/5/052012

Maries, A., & Singh, C. (2023). Helping Students Become Proficient Problem Solvers Part I: A Brief Review. *Education Sciences*, *13*(2), Article 2. https://doi.org/10.3390/educsci13020156

Medvedev, O. N., & Krägeloh, C. U. (2022). Rasch Measurement Model. In *Handbook of Assessment in Mindfulness Research* (pp. 1–18). Springer, Cham. https://doi.org/10.1007/978-3-030-77644-2_4-1

Mitchell, W. E., & Kowalik, T. F. (1979). Creative Problem Solving. *Gifted Child Quarterly*, *23*(3), 559–561. https://doi.org/10.1177/001698627902300322

Naingalis, M. L. P., Fauzi, A., Supurwoko, & Sartono, B. (2023). Item fit investigation of computational thinking scales instrument among Indonesian physics undergraduate students: A Rasch model analysis. *AIP Conference Proceedings*, *2540*(1), 090006. https://doi.org/10.1063/5.0105746

Park, Y.-J. (2021). Development of the 8-Item Phlegm Pattern Questionnaire (PPQ-8) Using Rasch Analysis. *Evidence-Based Complementary and Alternative Medicine*, *2021*. https://doi.org/10.1155/2021/6528891

Pears, M., Antoniou, P., Schiza, E., Matsangidou, M., Pattichis, C. S., Bamidis, P. D., & Konstantinidis, Stathis. Th. (2025). Enhancing creativity and cognitive skills in healthcare curricula: Recommendations from a modified delphi study on virtual reality integration. *Thinking Skills and Creativity*, *57*, 101810. https://doi.org/10.1016/j.tsc.2025.101810

Planinic, M., Boone, W. J., Susac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research: Why measurement matters. *Physical Review Physics Education Research*, *15*(2), 020111. https://doi.org/10.1103/PhysRevPhysEducRes.15.020111

Puccio, G. J., Klarman, B., & Szalay, P. A. (2022). Creative Problem-Solving. In *The Palgrave Encyclopedia of the Possible* (pp. 298–313). Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-90913-0_41

Purnami, W., Fauzi, A., & Naingalis, M. L. P. (2023). Computational thinking skills identification among students of physics education department using Rasch model analysis. *AIP Conference Proceedings*, *2751*(1), 100012. https://doi.org/10.1063/5.0143158

Ringo, S. S., Kuswanto, K., Samsudin, A., & Setiawan, A. (2021). Rural and urban students' attitudes toward physics: A comparative study using Rasch analysis. *Journal of Physics: Conference Series*, *1806*(1), 012009. https://doi.org/10.1088/1742-6596/1806/1/012009

Robertson, R. L., & Larki, M. J. (2019). Developing an Instrument to Observe and Evaluate Assessment System Maturity. *Journal of Educational Research and Practice*, *9*(1), 55–80. https://doi.org/10.5590/JERAP.2019.09.1.05

Shakhman, L., & Barak, M. (2019). The Physics Problem-Solving Taxonomy (PPST): Development and Application for Evaluating Student Learning. *EURASIA Journal of Mathematics, Science and Technology Education*, *15*(11). https://doi.org/10.29333/ejmste/109266

Shin, Y., Lee, H., & Kim, W. (2025). Exploring creative problem-solving in computer-supported collaborative learning: Focusing on group cohesiveness and socially shared metacognitive regulation. *Thinking Skills and Creativity*, *56*, 101771. https://doi.org/10.1016/j.tsc.2025.101771

Smith, G. T., & McCarthy, D. M. (1995). Methodological considerations in the refinement of clinical assessment instruments. *Psychological Assessment*, *7*(3), 300–308. https://doi.org/10.1037/1040-3590.7.3.300

Song, Y., Huang, L., Zheng, L., Fan, M., & Liu, Z. (2025). Interactions with generative AI chatbots: Unveiling dialogic dynamics, students' perceptions, and practical competencies in creative problem-solving. *International Journal of Educational Technology in Higher Education*, *22*(1). Scopus. https://doi.org/10.1186/s41239-025-00508-2

Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, *48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Testa, I., Capasso, G., Colantonio, A., Galano, S., Marzoli, I., Scotti di Uccio, U., & Serroni, G. (2020). Validation of University Entrance Tests Through Rasch Analysis. In M. S. Khine (Ed.), *Rasch Measurement: Applications in Quantitative Educational Research* (pp. 99–124). Springer. https://doi.org/10.1007/978-981-15-1800-3_6

Treffinger, D. J. (1995). Creative Problem Solving: Overview and educational implications. *Educational Psychology Review*, *7*(3), 301–312. https://doi.org/10.1007/BF02213375

Ulya, H., Sugiman, & Rosnawati, R. (2024). Design and content validity of mathematics creative problem-solving ability instrument for junior high school students. *Eurasia Journal of Mathematics, Science and Technology Education*, *20*(6), em2462. https://doi.org/10.29333/ejmste/14661

Viyanti, V., Rosidin, U., & Shintya, R. E. (2022). *Collaborative and Problem Solving Instruments in Project-Based Physics Learning | Indonesian Journal of Science and Mathematics Education*. https://doi.org/10.24042/ijsme.v5i1.11636

Widyaningsih, S. W., Yusuf, I., Prasetyo, Z. K., & Istiyono, E. (2021). The Development of the HOTS Test of Physics Based on Modern Test Theory: Question Modeling through E-Learning of Moodle LMS. *International Journal of Instruction*, *14*(4), 51–68. https://doi.org/10.29333/iji.2021.1444a

Xia, N., Haron, S. H., Huang, Y., & Niu, R. (2025). The effectiveness of CPS+SCAMPER teaching mode and strategies on student creativity. *Thinking Skills and Creativity*, *56*, 101758. https://doi.org/10.1016/j.tsc.2025.101758

Ying, Y., & Tiemann, R. (2024). Development of an assessment tool for collaborative problem-solving skills in chemistry. *Disciplinary and Interdisciplinary Science Education Research*, *6*(1), 25. https://doi.org/10.1186/s43031-024-00116-6

Yusuf, I., Widyaningsih, S. W., Prasetyo, Z. K., & Istiyono, E. (2020). Higher order thinking skills (HOTS)-oriented e-module in electric circuit. *Journal of Physics: Conference Series*, *1521*(2), 022027. https://doi.org/10.1088/1742-6596/1521/2/022027