



Instrumen Kemampuan *Creative Problem Solving* Siswa SMA pada Materi Energi Alternatif: Analisis Model Rasch

Icha Nurrohmah^{1*}, Lina Aviyanti¹, Ida Kaniawati¹

¹ Pendidikan Fisika, FPMIPA, Universitas Pendidikan Indonesia, Bandung, Indonesia.

DOI: <https://doi.org/10.29303/goescienceed.v7i2.1840>

Article Info:

Received : 25 April 2026
Revised : 30 April 2026
Accepted : 07 Mei 2026
Published : 11 Mei 2026

Correspondence:

Icha Nurrohmah

Phone: +6283845703868

Abstract: This study aims to evaluate the quality of an assessment instrument measuring students' Creative Problem Solving (CPS) skills in the context of alternative energy using the Rasch Model. CPS is a crucial competency in 21st-century physics learning; however, students' CPS abilities are often found to be relatively low, partly due to the limited availability of valid and reliable assessment instruments. This research employed a descriptive quantitative design involving 35 eleventh-grade students from a private senior high school in Bandung. The instrument consisted of 18 open-ended items analyzed using Ministep software based on the Rasch Model. The results indicated that the instrument satisfied the unidimensionality assumption, with a raw variance explained value of 51.2% and an unexplained variance in the first contrast of 8.9%, indicating that the instrument consistently measures a single CPS construct. All items demonstrated good model fit and high reliability, as reflected by a person reliability of 0.84, item reliability of 0.81, and Cronbach's Alpha of 0.93, suggesting that the instrument produces stable and reliable measurements. Furthermore, the distribution of item difficulty ranged from very easy to highly difficult, indicating that the instrument can effectively differentiate students' abilities across various proficiency levels. These findings confirm that the instrument is valid, reliable, and capable of providing more objective measurements of students' CPS skills. This study contributes to addressing the limitations of CPS assessment tools by providing a systematically validated instrument based on the Rasch Model, particularly in the context of alternative energy learning.

Keywords: Alternative Energy; Assessment Instrument; Creative Problem Solving Skills; Rasch Model; Physics Education.

Citation: Nurrohmah, I., Aviyanti, L., & Kaniawati, I. (2026). Instrumen Kemampuan Creative Problem Solving Siswa SMA pada Materi Energi Alternatif: Analisis Model Rasch. *Jurnal Pendidikan, Sains, Geologi, Dan Geofisika (GeoScienceEd Journal)*, 7(2), 1179-1186. <https://doi.org/10.29303/goescienceed.v7i2.1840>

Pendahuluan

Perkembangan ilmu pengetahuan dan teknologi yang terjadi secara pesat dewasa ini membawa dampak signifikan terhadap meningkatnya kebutuhan akan sumber daya manusia yang berkualitas dan adaptif dalam menghadapi tantangan global yang terus berubah. Pendidikan abad ke-21 menempatkan penguasaan empat keterampilan utama sebagai fondasi, yaitu *critical thinking*, *creativity*, *communication*, dan *collaboration*, yang tidak sekadar berorientasi pada

penguasaan konsep teoritis, melainkan juga menekankan kemampuan individu dalam mengatasi permasalahan kehidupan nyata secara efektif (Larson & Miller, 2011). Dalam ranah pembelajaran fisika, pengembangan keterampilan tersebut dapat difasilitasi melalui serangkaian kegiatan yang mendorong keterlibatan aktif peserta didik dalam proses berpikir mendalam dan pemecahan masalah secara kontekstual.

Kemampuan yang memerlukan perhatian lebih dalam proses pembelajaran fisika adalah kemampuan

Creative Problem Solving (CPS). Kemampuan ini didefinisikan sebagai kapabilitas individu dalam menciptakan gagasan-gagasan orisinal dan solusi inovatif melalui tahapan yang terstruktur, mencakup *objective finding*, *fact finding*, *problem finding*, *idea finding*, *solution finding*, dan *acceptance finding* (Treffinger & Isaksen, 2005). *CPS* memadukan proses kreatif dengan keterampilan pemecahan masalah, sehingga peserta didik memperoleh bekal untuk menghadapi berbagai persoalan kompleks secara rasional dan inovatif. (Hu, Xiaohui, & Shieh, 2017) menegaskan bahwa kemampuan *CPS* bukan hanya dibutuhkan dalam konteks akademik, melainkan juga berperan penting dalam kehidupan sehari-hari untuk meraih keberhasilan dan keunggulan kompetitif.

Topik energi alternatif merupakan salah satu materi yang dinilai strategis untuk menumbuhkan kemampuan *CPS* peserta didik. Hal ini dilandasi oleh sifat materi tersebut yang bersifat lintas disiplin dan berkaitan erat dengan persoalan-persoalan global kontemporer, seperti krisis energi dan ancaman terhadap keberlanjutan lingkungan hidup. Pembahasan seputar energi alternatif mendorong peserta didik untuk berpikir secara terbuka, mengeksplorasi beragam alternatif solusi, serta menilai kelayakannya baik dari perspektif ilmiah maupun sosial (Dinantika, Suyanto, & Nyeneng, 2019).

Sejumlah hasil penelitian mengungkapkan bahwa kemampuan *CPS* peserta didik secara umum masih tergolong dalam kategori sedang hingga rendah (Junita, 2016, 2016; Khalid et al., 2020; Nuraziza & Suwarma, 2018). Kondisi serupa turut teridentifikasi dalam studi pendahuluan yang dilakukan di salah satu SMA swasta di Kota Bandung. Data menunjukkan bahwa capaian peserta didik pada masing-masing indikator *CPS* masih berada di bawah harapan, dengan persentase *objective finding* 52,63%, *fact finding* 45,79%, *problem finding* 45,79%, *idea finding* 41,05%, *solution finding* 41,05%, dan *acceptance finding* 48,42%. Rata-rata capaian keseluruhan hanya sebesar 46,32%, yang terklasifikasi dalam kategori sangat kurang kreatif. Kondisi ini mengindikasikan bahwa kemampuan *CPS* belum terlatih secara maksimal dalam kegiatan pembelajaran, sehingga dibutuhkan instrumen yang tepat guna mengukur dan menganalisis kemampuan tersebut secara lebih mendalam.

Rendahnya capaian *CPS* peserta didik tidak semata-mata disebabkan oleh faktor proses pembelajaran, tetapi juga dipengaruhi oleh keterbatasan kualitas instrumen evaluasi yang digunakan. Instrumen berkualitas tinggi harus memenuhi standar validitas dan reliabilitas agar mampu menghasilkan gambaran yang sah mengenai kemampuan peserta didik. Pendekatan *Classical Test Theory (CTT)* diketahui memiliki kelemahan inheren karena bersifat *sample-dependent* dan tidak mampu memisahkan secara objektif

antara kemampuan peserta didik dan tingkat kesukaran butir soal (Sumintono & Widhiarso, 2015).

Model *Rasch* yang merupakan bagian dari kerangka *Item Response Theory (IRT)* menawarkan analisis yang lebih objektif dan presisi. Model ini mampu mengestimasi kemampuan peserta didik (*person ability*) secara terpisah dari tingkat kesukaran butir soal (*item difficulty*), serta menghasilkan informasi yang lebih komprehensif terkait validitas, reliabilitas, dan kesesuaian butir soal (Sumintono & Widhiarso, 2015). Berbagai penelitian telah membuktikan efektivitas Efektivitas Model *Rasch* dalam mengevaluasi kualitas instrumen penilaian pendidikan telah dibuktikan oleh berbagai studi terdahulu (Khotimah, Hadi, & Lestari, 2024; Neva Lionitha Ibrahim, 2025, 2025; Permatasari, 2025; Prieto & Delgado, 2003).

Kajian yang secara spesifik menganalisis kualitas instrumen kemampuan *CPS* pada konteks materi energi alternatif masih sangat minim. Mayoritas penelitian yang ada lebih menitikberatkan pada strategi peningkatan kemampuan *CPS* melalui model pembelajaran tertentu, tanpa disertai evaluasi mendalam terhadap kualitas instrumen yang digunakan.

Penelitian ini difokuskan pada pengembangan dan analisis komprehensif instrumen kemampuan *CPS* pada materi energi alternatif menggunakan Model *Rasch*, mencakup dimensi validitas, reliabilitas, dan tingkat kesukaran butir soal. Instrumen yang dihasilkan diharapkan mampu menjadi alat ukur yang lebih objektif dan dapat diandalkan dalam mengevaluasi kemampuan *Creative Problem Solving* peserta

Metode

Penelitian ini dirancang sebagai studi kuantitatif deskriptif yang bertujuan untuk memaparkan kualitas instrumen pengukuran kemampuan *Creative Problem Solving* pada topik energi alternatif. Pendekatan ini dipilih karena penelitian menitikberatkan pada pengkajian kualitas instrumen secara sistematis berdasarkan data empiris yang bersumber dari hasil uji coba langsung pada peserta didik. Subjek penelitian adalah peserta didik kelas XI di salah satu SMA swasta yang berlokasi di Kota Bandung. Penentuan sampel dilakukan melalui teknik *purposive sampling* dengan mempertimbangkan relevansi materi energi alternatif terhadap jenjang kelas XI, sehingga diperoleh total sampel sebanyak 35 peserta didik.

Instrumen yang digunakan dalam penelitian ini berupa tes kemampuan *Creative Problem-Solving* dalam format soal uraian yang terdiri dari 18 butir. Pengembangan instrumen dilakukan dengan mengadaptasi perangkat dari penelitian sebelumnya, yang selanjutnya disesuaikan dengan karakteristik materi energi alternatif dan kondisi peserta didik.

Penyusunan butir soal mengacu pada enam indikator CPS yang dikemukakan oleh Treffinger dan Isaksen (2005), meliputi: (1) *objective finding*, (2) *fact finding*, (3) *problem finding*, (4) *idea finding*, (5) *solution finding*, dan (6) *acceptance finding*, dengan setiap indikator direpresentasikan oleh tiga butir soal.

Pelaksanaan penelitian menempuh lima tahapan. Tahap pertama berupa studi pendahuluan yang dilakukan melalui pemberian tes awal (*preliminary test*) kemampuan *Creative Problem Solving* kepada peserta didik di salah satu SMA swasta di Kota Bandung. Data hasil tes ini digunakan untuk memperoleh gambaran awal kemampuan CPS peserta didik sebagai dasar dalam merumuskan masalah penelitian. Kedua, penyusunan instrumen berdasarkan indikator CPS yang dikontekstualisasikan pada materi energi alternatif. Ketiga, validasi ahli (*expert judgment*) yang melibatkan dosen ahli untuk menilai kesesuaian aspek isi, konstruk, dan kebahasaan instrumen, dilanjutkan dengan revisi berdasarkan saran yang diperoleh. Keempat, pelaksanaan uji coba instrumen kepada 35 peserta didik kelas XI. Kelima, analisis data menggunakan Model *Rasch* dengan dukungan perangkat lunak *Ministep*.

Data yang diperoleh dari uji coba dianalisis menggunakan Model *Rasch* karena model ini memiliki sejumlah kelebihan dibandingkan *Classical Test Theory (CTT)*, antara lain bersifat *person free* dan *item free*, mampu memetakan distribusi kemampuan peserta didik sekaligus tingkat kesukaran butir soal, dapat mengakomodasi data yang tidak lengkap, serta menghasilkan informasi yang lebih objektif dan komprehensif mengenai kualitas instrumen (Perdana, 2018; Sumintono & Widhiarso, 2015).

Analisis data dilaksanakan dalam tiga aspek utama. Pertama, analisis validitas instrumen dilakukan melalui pengujian unidimensionalitas berdasarkan nilai *raw variance explained by measure* dan *unexplained variance in 1st contrast*, serta penilaian kelayakan butir soal menggunakan parameter *Outfit Mean Square (MNSQ)*, *Outfit Z-Standard (ZSTD)*, dan *Point Measure Correlation (Pt Measure Corr.)*. Kedua, analisis reliabilitas mencakup nilai *person reliability*, *item reliability*, dan *Cronbach's Alpha*. Ketiga, analisis tingkat kesukaran butir soal dilakukan berdasarkan nilai logit pada tabel *Item Measure* serta peta visualisasi *Wright Map*. Seluruh hasil analisis kemudian diinterpretasikan merujuk pada kriteria yang ditetapkan oleh (Sumintono & Widhiarso, 2015).

Hasil dan Diskusi

Analisis Validitas instrumen

Unidimensionalitas merupakan salah satu indikator krusial dalam evaluasi instrumen yang mencerminkan kemampuan alat ukur untuk mengukur satu konstruk tunggal secara konsisten dan terarah.

Dalam penelitian ini, pengujian unidimensionalitas dilaksanakan dengan memanfaatkan perangkat lunak *MINISTEP Rasch* melalui penelaahan *output Item Dimensionality*. Hasil pengujian selanjutnya diinterpretasikan berdasarkan kriteria yang tercantum pada Tabel 1.

Tabel 1. Acuan Interpretasi Nilai Unidimensionalitas Instrumen

| Nilai Raw Variance Explained by Measure (%) | Kriteria |
|---|-----------|
| 60 < Rve ≤ 100 | Istimewa |
| 40 < Rve ≤ 60 | Sesuai |
| 20 ≤ Rve ≤ 40 | Terpenuhi |

(Sumintono & Widhiarso, 2015)

Nilai *raw variance explained by measure*, unidimensionalitas instrumen turut dapat dikaji melalui nilai *unexplained variance in 1st contrast*. Nilai ini berfungsi untuk mendeteksi kemungkinan adanya dimensi tambahan yang muncul di luar konstruk utama dalam instrumen. Acuan penilaian untuk parameter tersebut disajikan pada Tabel 2.

Tabel 2. Acuan Penilaian Nilai Unexplained Variance in 1st Contrast

| Nilai Unexplained Variance (%) | Kriteria |
|--------------------------------|--------------|
| < 15 | Buruk |
| 10 - 15 | Cukup |
| 5 - 10 | Bagus |
| 3 - 5 | Sangat Bagus |
| > 3 | Istimewa |

(Sumintono & Widhiarso, 2015)

Hasil pengujian unidimensionalitas instrumen kemampuan *Creative Problem Solving* yang diperoleh dari *output tables 23: Item Dimensionality* pada perangkat lunak *MINISTEP Rasch* ditampilkan pada Gambar 1.

| | Eigenvalue | Observed | Expected |
|--|------------|----------|----------|
| Total raw variance in observations = | 36.9181 | 100.0% | 100.0% |
| Raw variance explained by measures = | 18.9181 | 51.2% | 52.8% |
| Raw variance explained by persons = | 13.1147 | 35.5% | 36.6% |
| Raw Variance explained by items = | 5.8034 | 15.7% | 16.2% |
| Raw unexplained variance (total) = | 18.0000 | 48.8% | 47.2% |
| Unexplained variance in 1st contrast = | 3.3025 | 8.9% | 18.3% |
| Unexplained variance in 2nd contrast = | 2.5524 | 6.9% | 14.2% |
| Unexplained variance in 3rd contrast = | 1.6489 | 4.5% | 9.2% |

Gambar 1. Hasil Uji Unidimensionalitas Instrumen Kemampuan *Creative Problem Solving*

Hasil pengujian mengungkapkan bahwa nilai *raw variance explained by measure* yang dihasilkan dari uji coba instrumen adalah sebesar 51,2%. Perolehan ini mengindikasikan bahwa lebih dari setengah variansi data yang diamati mampu dijelaskan oleh instrumen, sehingga instrumen dinilai cukup representatif dalam menangkap konstruk yang hendak diukur. Sementara

itu, nilai *unexplained variance in 1st contrast* sebesar 8,9% masih berada di bawah ambang batas yang dipersyaratkan. Hal ini mengonfirmasi bahwa instrumen tidak terpengaruh secara substansial oleh dimensi-dimensi di luar konstruk utama yang diukur, sehingga asumsi unidimensionalitas Model *Rasch* dapat terpenuhi dengan baik.

Validitas instrumen secara keseluruhan terkonfirmasi, analisis dilanjutkan pada level butir soal dengan menggunakan tiga parameter utama, yaitu *Outfit Mean Square* (MNSQ), *Outfit Z-Standard* (ZSTD), dan *Point Measure Correlation* (Pt Measure Corr.). Ketiga parameter ini diterapkan secara simultan untuk mengevaluasi ketepatan setiap butir soal terhadap model *Rasch*. Kriteria kelayakan butir soal yang menjadi acuan penilaian disajikan pada Tabel 3.

Tabel 3. Acuan Interpretasi Nilai MNSQ dan ZSTD

| Outfit | Kriteria |
|------------------|-------------------------------|
| MNSQ | 0,5 < MNSQ < 1,5 |
| ZSTD | -2,0 < ZSTD < +2,0 |
| Pt Measure Corr. | 0,4 < Pt Measure Corr. < 0,85 |

(Sumintono & Widhiarso, 2015).

Hasil dari ketiga parameter tersebut kemudian dikategorikan berdasarkan acuan penilaian kualitas butir soal yang tercantum pada Tabel 4.

Tabel 4. Acuan Penilaian Kualitas Butir Soal Berdasarkan *Fit-Statistic*

| Pemenuhan Kriteria Nilai | Kategori |
|------------------------------------|---------------|
| Seluruh kriteria terpenuhi | Sangat sesuai |
| Dua dari tiga kriteria terpenuhi | Sesuai |
| Hanya satu dari kriteria terpenuhi | Kurang Sesuai |
| Tidak ada kriteria terpenuhi | Tidak Sesuai |

(Sumintono & Widhiarso, 2015).

Tabel 5. Ringkasan Hasil Uji Validitas Setiap Butir Soal

| Butir Soal | Nilai Outfit | | Pt. Measure Corr. | Pemenuhan Kriteria | Kategori |
|------------|--------------|--------|-------------------|----------------------------------|---------------|
| | MNSQ | ZSTD | | | |
| 1a | 1,09 | 0,35 | 0,56 | Seluruh kriteria terpenuhi | Sangat Sesuai |
| 1b | 0,64 | - 1,61 | 0,74 | Seluruh kriteria terpenuhi | Sangat Sesuai |
| 1c | 1,54 | 1,28 | 0,51 | Dua dari tiga kriteria terpenuhi | Sesuai |
| 1d | 0,59 | -0,95 | 0,65 | Seluruh kriteria terpenuhi | Sangat Sesuai |
| 1e | 1,03 | 0,20 | 0,63 | Seluruh kriteria terpenuhi | Sangat Sesuai |
| 1f | 1,91 | 1,88 | 0,46 | Dua dari tiga kriteria terpenuhi | Sesuai |
| 2a | 1,09 | 0,35 | 0,61 | Seluruh kriteria terpenuhi | Sangat Sesuai |
| 2b | 0,86 | -0,49 | 0,70 | Seluruh kriteria terpenuhi | Sangat Sesuai |
| 2c | 0,80 | -0,44 | 0,64 | Seluruh kriteria terpenuhi | Sangat Sesuai |
| 2d | 0,96 | -0,02 | 0,70 | Seluruh kriteria terpenuhi | Sangat Sesuai |
| 2e | 0,79 | 0,59 | 0,64 | Seluruh kriteria terpenuhi | Sangat Sesuai |

Hasil analisis kesesuaian setiap butir soal terhadap model diperoleh melalui *output* tabel 10: *Item Fit Order Dimensionality* pada perangkat lunak *MINISTEP Rasch* sebagaimana ditampilkan pada Gambar 2.

| ENTRY NUMBER | TOTAL SCORE | TOTAL COUNT | JMLE MEASURE | MODEL S.E. | INFIT MNSQ | INFIT ZSTD | OUTFIT MNSQ | OUTFIT ZSTD | PTMEASURE CORR. | EXACT MATCH EXP. | EXACT MATCH OBSV. | EXACT MATCH EXP% | ITEM |
|--------------|-------------|-------------|--------------|------------|------------|------------|-------------|-------------|-----------------|------------------|-------------------|------------------|------|
| 2 | 111 | 34 | 1.41 | .20 | .66 | -1.56 | .64 | -1.61 | .74 | .71 | 54.5 | 40.2 | 1b |
| 14 | 117 | 34 | 1.17 | .20 | .99 | .01 | .93 | -.22 | .67 | .70 | 33.3 | 42.5 | 3b |
| 8 | 121 | 34 | 1.01 | .20 | .93 | -.21 | .86 | -.49 | .70 | .69 | 48.5 | 43.3 | 2b |
| 10 | 139 | 34 | .22 | .22 | .60 | -1.69 | .96 | -.02 | .70 | .64 | 66.7 | 50.9 | 2d |
| 11 | 140 | 34 | .17 | .23 | .96 | -.05 | .79 | -.59 | .64 | .63 | 42.4 | 51.1 | 2e |
| 12 | 141 | 34 | .11 | .23 | 1.45 | 1.57 | 1.49 | 1.39 | .49 | .63 | 45.5 | 52.3 | 2f |
| 16 | 142 | 34 | .06 | .23 | .86 | -.47 | .70 | -.87 | .68 | .63 | 54.5 | 52.6 | 3d |
| 5 | 146 | 34 | -.17 | .24 | 1.20 | .75 | 1.03 | .20 | .63 | .60 | 54.5 | 57.4 | 1e |
| 1 | 147 | 34 | -.23 | .25 | 1.28 | 1.27 | 1.09 | .35 | .56 | .60 | 60.6 | 57.8 | 1a |
| 7 | 147 | 34 | -.23 | .25 | 1.28 | 1.00 | 1.09 | .35 | .61 | .60 | 54.5 | 57.8 | 2a |
| 9 | 147 | 34 | -.23 | .25 | .81 | -.63 | .80 | -.44 | .64 | .60 | 57.6 | 57.8 | 2c |
| 3 | 149 | 34 | -.35 | .26 | 1.35 | 1.16 | 1.54 | 1.28 | .51 | .59 | 63.6 | 59.5 | 1c |
| 15 | 149 | 34 | -.35 | .26 | .73 | -.90 | .64 | -.89 | .66 | .59 | 54.5 | 59.5 | 3c |
| 17 | 149 | 34 | -.35 | .26 | 1.19 | .70 | .95 | .00 | .58 | .59 | 51.5 | 59.5 | 3e |
| 6 | 150 | 34 | -.42 | .26 | 1.63 | 1.86 | 1.91 | 1.88 | .46 | .58 | 60.6 | 60.8 | 1f |
| 4 | 152 | 34 | -.56 | .27 | .65 | -1.16 | .59 | -.95 | .65 | .56 | 72.7 | 64.8 | 1d |
| 13 | 152 | 34 | -.56 | .27 | 1.25 | .86 | .87 | -.16 | .60 | .56 | 63.6 | 64.8 | 3a |
| 18 | 154 | 34 | -.71 | .28 | 1.00 | -.12 | .80 | -.31 | .57 | .55 | 51.5 | 66.3 | 3f |
| MEAN | 141.8 | 34.0 | .00 | .24 | 1.05 | -.15 | .98 | -.06 | | | 55.1 | 55.5 | |
| P.SD | 12.2 | .0 | .59 | .02 | .29 | 1.04 | .34 | .86 | | | 9.0 | 7.4 | |

Gambar 2. Hasil Analisis Validitas Instrumen Kemampuan *Creative Problem Solving*

Berdasarkan hasil analisis tersebut, rangkuman tingkat kesesuaian setiap butir soal pada instrumen kemampuan *Creative Problem Solving* disajikan secara lengkap pada Tabel 5.

Merujuk pada Tabel 5, hasil evaluasi kesesuaian butir soal menunjukkan bahwa sebanyak 16 butir soal berhasil memenuhi ketiga parameter kelayakan secara bersamaan sehingga dikategorikan sangat sesuai, sementara 2 butir soal lainnya hanya memenuhi dua dari tiga parameter yang dipersyaratkan sehingga masuk dalam kategori sesuai. Berdasarkan temuan ini, dapat ditegaskan bahwa seluruh 18 butir soal dinyatakan *fit* terhadap Model *Rasch* dan layak digunakan sebagai instrumen pengukuran kemampuan *CPS* peserta didik.

| Butir Soal | Nilai Outfit | | Pt. Measure Corr. | Pemenuhan Kriteria | Kategori |
|------------|--------------|-------|-------------------|----------------------------|---------------|
| | MNSQ | ZSTD | | | |
| 2f | 1,49 | 1,39 | 0,49 | Seluruh kriteria terpenuhi | Sangat Sesuai |
| 3a | 0,87 | 0,16 | 0,60 | Seluruh kriteria terpenuhi | Sangat Sesuai |
| 3b | 0,93 | -0,22 | 0,67 | Seluruh kriteria terpenuhi | Sangat Sesuai |
| 3c | 0,64 | 0,89 | 0,66 | Seluruh kriteria terpenuhi | Sangat Sesuai |
| 3d | 0,70 | -0,87 | 0,68 | Seluruh kriteria terpenuhi | Sangat Sesuai |
| 3e | 0,95 | 0,00 | 0,58 | Seluruh kriteria terpenuhi | Sangat Sesuai |
| 3f | 0,80 | -0,31 | 0,57 | Seluruh kriteria terpenuhi | Sangat Sesuai |

Berdasarkan serangkaian analisis menggunakan Model Rasch, instrumen kemampuan *Creative Problem Solving* pada materi energi alternatif terbukti memiliki kualitas yang baik ditinjau dari aspek validitas, reliabilitas, maupun sebaran tingkat kesukaran butir soal. Pengujian unidimensionalitas menghasilkan nilai *raw variance explained by measure* sebesar 51,2% yang masuk dalam kategori "sesuai", menandakan bahwa instrumen mampu mengukur konstruk yang dimaksud secara memadai. Di samping itu, nilai *unexplained variance in 1st contrast* sebesar 8,9% berada pada kategori "bagus", yang mengindikasikan bahwa instrumen tidak dipengaruhi secara berarti oleh dimensi-dimensi lain di luar konstruk utama. Kondisi ini membuktikan bahwa instrumen telah memenuhi prasyarat dasar Model Rasch, yakni pengukuran satu konstruk secara konsisten (Sumintono & Widhiarso, 2015).

Dari sisi validitas butir soal, seluruh butir berada pada rentang kategori "sesuai" hingga "sangat sesuai", yang mengindikasikan bahwa instrumen mampu merepresentasikan konstruk kemampuan *Creative Problem Solving* dengan baik. Hasil ini selaras dengan temuan penelitian terdahulu yang menyatakan bahwa Model Rasch mampu memberikan penilaian kualitas butir soal yang lebih objektif dibandingkan pendekatan konvensional (Neva Lionitha Ibrahim, 2025). Dengan demikian, setiap butir dalam instrumen ini dapat diandalkan untuk mengukur kemampuan peserta didik secara akurat.

Analisis Reliabilitas

Nilai *Alpha Cronbach* berperan sebagai indikator yang menggambarkan sejauh mana keterkaitan antara responden dengan keseluruhan butir soal yang termuat dalam instrumen. Panduan interpretasi nilai tersebut secara lengkap tersaji pada Tabel 6.

Tabel 5. Acuan Interpretasi Nilai *Alpha Cronbach*

| Rentang nilai | Kriteria |
|--------------------------|--------------|
| $KR - 20 < 0,5$ | Buruk |
| $0,5 \leq KR - 20 < 0,6$ | Jelek |
| $0,6 \leq KR - 20 < 0,7$ | Cukup |
| $0,7 \leq KR - 20 < 0,8$ | Bagus |
| $KR - 20 < 0,8$ | Sangat Bagus |

(Sumintono & Widhiarso, 2015)

Nilai *person reliability* mencerminkan konsistensi peserta didik dalam merespons seluruh butir soal yang ada, sedangkan nilai *item reliability* menggambarkan kestabilan butir-butir soal dalam instrumen secara keseluruhan. Kedua nilai ini diekstraksi dari *output Tabel 3.1 Summary Statistics* pada perangkat lunak MINISTEP Rasch. Panduan interpretasi untuk kedua nilai reliabilitas tersebut disajikan pada Tabel 7.

Tabel 6. Acuan Interpretasi Nilai Person Reliability dan Item Reliability

| Skala | Kriteria |
|----------------------|--------------|
| $r \leq 0,67$ | Lemah |
| $0,67 < r \leq 0,80$ | Cukup |
| $0,80 < r \leq 0,90$ | Bagus |
| $0,90 < r \leq 0,94$ | Bagus Sekali |
| $r < 0,94$ | Istimewa |

(Sumintono & Widhiarso, 2015).

Hasil uji reliabilitas diekstrak dari *output tables: 3.1 Summary Statistic* pada perangkat lunak MINISTEP Rasch sebagaimana ditampilkan pada Gambar 3 dan Gambar 4.

The image shows two screenshots of Rasch output tables. The top screenshot is titled 'SUMMARY OF 33 MEASURED (NON-EXTREME) PERSON' and includes a table with columns for Total Score, Count, Measure, Model S.E., Infit MNSQ, ZSTD, and Outfit MNSQ, ZSTD. Below this table, it lists 'PERSON RELIABILITY .86' and 'PERSON RELIABILITY .88'. The bottom screenshot is titled 'SUMMARY OF 34 MEASURED (EXTREME AND NON-EXTREME) PERSON' and includes a similar table. Below it, it lists 'PERSON RELIABILITY .84' and 'PERSON RELIABILITY .85'. At the bottom of the second screenshot, it shows 'CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .93' and 'STANDARDIZED (50 ITEM) RELIABILITY = .94'.

Gambar 3. Hasil analisis *Person reliability*

| SUMMARY OF 18 MEASURED (NON-EXTREME) ITEM | | | | | | | | |
|---|-------------|---------|---------|------------|------------|------------------|-------------|-------|
| | TOTAL SCORE | COUNT | MEASURE | MODEL S.E. | INFIT MNSQ | ZSTD | OUTFIT MNSQ | ZSTD |
| MEAN | 141.8 | 34.0 | .00 | .24 | 1.05 | .15 | .98 | -.06 |
| SEM | 3.0 | .0 | .14 | .01 | .07 | .25 | .08 | .21 |
| P.SD | 12.2 | .0 | .59 | .02 | .29 | 1.04 | .34 | .86 |
| S.SD | 12.6 | .0 | .61 | .03 | .30 | 1.07 | .35 | .89 |
| MAX. | 154.0 | 34.0 | 1.41 | .28 | 1.63 | 1.86 | 1.91 | 1.88 |
| MIN. | 111.0 | 34.0 | -.71 | .20 | .60 | -1.69 | .59 | -1.61 |
| REAL RMSE | .26 | TRUE SD | .53 | SEPARATION | 2.03 | ITEM RELIABILITY | .81 | |
| MODEL RMSE | .24 | TRUE SD | .54 | SEPARATION | 2.23 | ITEM RELIABILITY | .83 | |
| S.E. OF ITEM MEAN = .14 | | | | | | | | |
| ITEM RAW SCORE-TO-MEASURE CORRELATION = -1.00 | | | | | | | | |
| Global statistics: please see Table 44. | | | | | | | | |
| UMEAN=.0000 USCALE=1.0000 | | | | | | | | |

Gambar 4. Hasil analisis Item Reliability

Berdasarkan Gambar 3, nilai *person reliability* yang diperoleh adalah sebesar 0,84 yang tergolong dalam kategori bagus. Pada Gambar 4, nilai *item reliability* tercatat sebesar 0,81 dan juga berada dalam kategori bagus. Adapun nilai *Alpha Cronbach* yang dihasilkan sebesar 0,93 termasuk dalam kategori sangat bagus. Berdasarkan keseluruhan perolehan tersebut, dapat disimpulkan bahwa instrumen tes kemampuan *Creative Problem-Solving* yang dikembangkan telah memenuhi standar reliabilitas yang dipersyaratkan dan layak digunakan dalam kegiatan penelitian.

Ditinjau dari aspek reliabilitas, instrumen menunjukkan performa yang tinggi dengan nilai *person reliability* 0,84 dan *item reliability* 0,81, keduanya berada dalam kategori "bagus". Nilai *Alpha Cronbach* sebesar 0,93 pun termasuk dalam kategori "sangat bagus", mencerminkan konsistensi internal instrumen yang sangat kuat. Hal ini menandakan bahwa instrumen mampu menghasilkan data pengukuran yang stabil dan dapat diandalkan. Temuan ini memperkuat hasil penelitian sebelumnya yang menunjukkan bahwa penerapan Model *Rasch* berkontribusi pada peningkatan kualitas reliabilitas instrumen dalam penelitian pendidikan (Marjiastuti & Wahyuni, 2015; Neva Lionitha Ibrahim, 2025; Permatasari, 2025).

Analisis Tingkat Kesukaran Butir Soal

Tingkat kesukaran masing-masing butir soal dianalisis dengan mengacu pada dua keluaran utama perangkat lunak *MINISTEP Rasch*, yakni *output tabel 13: Item Measure* dan *output tabel 1: Variable (Wright Maps)*.

Tabel 7. Acuan Kesukaran butir soal

| Tingkat kesukaran | Kategori |
|------------------------|--------------|
| $ME \leq -1 SD$ | Sangat Mudah |
| $-1 SD < ME \leq 0,00$ | Mudah |
| $0,00 < ME \leq +1 SD$ | Sukar |
| $+1 SD < ME$ | Sangat Sukar |

(Sumintono & Widhiarso, 2015)

Tabel 13 digunakan untuk mengklasifikasikan setiap butir soal ke dalam kategori tingkat kesukaran tertentu berdasarkan nilai logit dan standar deviasi yang

dihasilkan, sedangkan tabel 1 dimanfaatkan untuk memperoleh gambaran menyeluruh tentang pola respons peserta didik terhadap instrumen sekaligus memetakan kekuatan relatif setiap butir soal. Berdasarkan hasil analisis tingkat kesukaran masing-masing butir soal ditampilkan pada Gambar 5 berikut.

| ITEM STATISTICS: MEASURE ORDER | | | | | | | | | | | | | |
|--------------------------------|-------------|-------------|-------------|------------|------------|-------|-------------|-------|-------------------|------------|------------|------|----|
| ENTRY NUMBER | TOTAL SCORE | TOTAL COUNT | MLE MEASURE | MODEL S.E. | INFIT MNSQ | ZSTD | OUTFIT MNSQ | ZSTD | PTMEASUR-AL CORR. | EXACT EXP. | MATCH OBSK | ITEM | |
| 2 | 111 | 34 | 1.41 | .20 | .66 | -1.56 | .64 | -1.61 | .74 | .71 | 54.5 | 40.2 | 1b |
| 14 | 117 | 34 | 1.17 | .20 | .99 | .01 | .93 | -.22 | .67 | .70 | 33.3 | 42.5 | 3b |
| 8 | 121 | 34 | 1.01 | .20 | .93 | -.21 | .86 | -.49 | .70 | .69 | 48.5 | 43.3 | 2b |
| 10 | 139 | 34 | .22 | .22 | .60 | -1.69 | .96 | -.02 | .70 | .64 | 66.7 | 50.9 | 2d |
| 11 | 140 | 34 | .17 | .23 | .96 | -.05 | .79 | -.59 | .64 | .63 | 42.4 | 51.1 | 2e |
| 12 | 141 | 34 | -.11 | .23 | 1.45 | 1.57 | 1.49 | 1.39 | .49 | .63 | 45.5 | 52.3 | 2f |
| 16 | 142 | 34 | .06 | .23 | .86 | -.47 | .70 | -.87 | .68 | .63 | 54.5 | 52.6 | 3d |
| 5 | 146 | 34 | -.17 | .24 | 1.20 | .75 | 1.03 | .20 | .63 | .60 | 54.5 | 57.4 | 1e |
| 1 | 147 | 34 | -.23 | .25 | 1.38 | 1.27 | 1.09 | .35 | .56 | .60 | 60.6 | 57.8 | 1a |
| 7 | 147 | 34 | -.23 | .25 | 1.28 | 1.00 | 1.09 | .35 | .61 | .60 | 54.5 | 57.8 | 2a |
| 9 | 147 | 34 | -.23 | .25 | .81 | -.63 | .80 | -.44 | .64 | .60 | 57.6 | 57.8 | 2c |
| 3 | 149 | 34 | -.35 | .26 | 1.35 | 1.16 | 1.54 | 1.28 | .51 | .59 | 63.6 | 59.5 | 1c |
| 15 | 149 | 34 | -.35 | .26 | .73 | -.90 | .64 | -.89 | .66 | .59 | 54.5 | 59.5 | 3c |
| 17 | 149 | 34 | -.35 | .26 | 1.19 | .70 | .95 | .00 | .58 | .59 | 51.5 | 59.5 | 3e |
| 6 | 150 | 34 | -.42 | .26 | 1.63 | 1.86 | 1.91 | 1.88 | .46 | .58 | 60.6 | 60.8 | 1f |
| 4 | 152 | 34 | -.56 | .27 | .65 | -1.16 | .59 | -.95 | .65 | .56 | 72.7 | 64.8 | 1d |
| 13 | 152 | 34 | -.56 | .27 | 1.25 | .86 | .87 | -.16 | .60 | .56 | 63.6 | 64.8 | 3a |
| 18 | 154 | 34 | -.71 | .28 | 1.00 | .12 | .80 | -.31 | .57 | .55 | 51.5 | 66.3 | 3f |
| MEAN | 141.8 | 34.0 | .00 | .24 | 1.05 | .15 | .98 | -.06 | | | 55.1 | 55.5 | |
| P.SD | 12.2 | .0 | .59 | .02 | .29 | 1.04 | .34 | .86 | | | 9.0 | 7.4 | |

Gambar 5. Hasil Tingkat Kesukaran Berdasarkan Output item Measure

Berdasarkan hasil analisis tersebut, klasifikasi tingkat kesukaran masing-masing butir soal disajikan secara rinci pada Tabel 9.

Tabel 8. Klasifikasi Tingkat Kesukaran Setiap Butir Soal

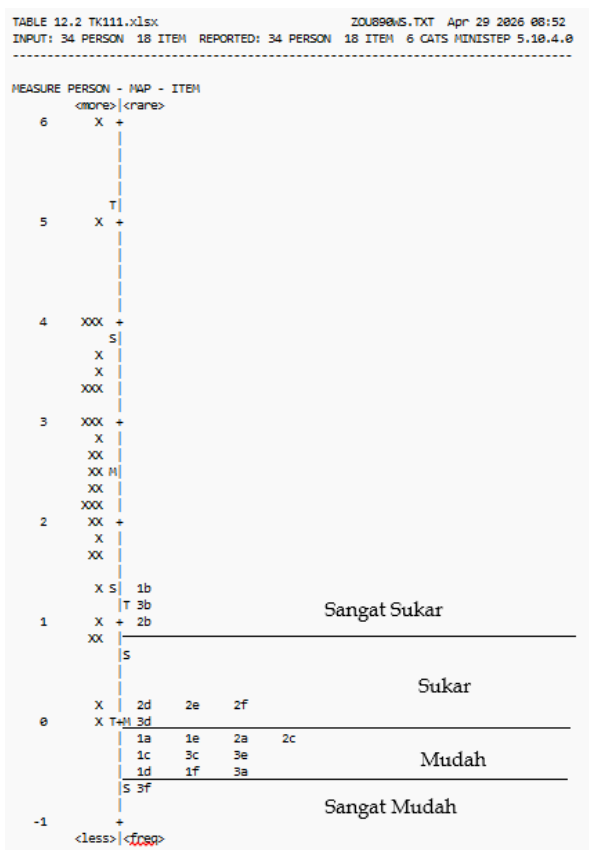
| Butir Soal | Nilai Measure | Kategori |
|------------|---------------|--------------|
| 1a | -0,23 | Mudah |
| 1b | 1,41 | Sangat sukar |
| 1c | -0,35 | Mudah |
| 1d | -0,56 | Mudah |
| 1e | -0,17 | Mudah |
| 1f | -0,42 | Mudah |
| 2a | -0,23 | Mudah |
| 2b | 1,01 | Sangat sukar |
| 2c | -0,23 | Mudah |
| 2d | 0,22 | Sukar |
| 2e | 0,17 | Sukar |
| 2f | 0,11 | Sukar |
| 3a | -0,56 | Mudah |
| 3b | 1,17 | Sangat sukar |
| 3c | 0,35 | Sukar |
| 3d | 0,06 | Sukar |
| 3e | -0,35 | Mudah |
| 3f | -0,71 | Sangat Mudah |

Merujuk pada Tabel 9, sebaran tingkat kesukaran butir soal memperlihatkan bahwa satu butir soal tergolong sangat mudah, sepuluh butir soal tergolong mudah, empat butir soal tergolong sukar, dan tiga butir soal tergolong sangat sukar. Pola sebaran ini mencerminkan bahwa instrumen memiliki variasi

tingkat kesukaran yang proporsional, sehingga mampu mengakomodasi pengukuran kemampuan *Creative Problem-Solving* peserta didik pada berbagai tingkatan secara lebih menyeluruh dan komprehensif.

Analisis tingkat kesukaran butir soal memperlihatkan adanya variasi yang mencakup kategori sangat mudah, mudah, sukar, hingga sangat sukar. Keberagaman tingkat kesukaran ini merupakan hal yang penting dalam sebuah instrumen karena memungkinkan pengukuran kemampuan peserta didik secara lebih menyeluruh. Butir soal dengan tingkat kesukaran yang bervariasi secara proporsional diperlukan untuk membedakan kemampuan peserta didik pada berbagai level. Dengan sebaran tingkat kesukaran yang seimbang, instrumen ini dinilai mampu mengidentifikasi kemampuan *Creative Problem Solving* peserta didik dengan lebih akurat dan komprehensif.

Sebaran tingkat kesukaran pada setiap indikator kemampuan *Creative Problem Solving* peserta didik berdasarkan *Wright Map* dapat dilihat pada gambar berikut.



Gambar 6. *Wright Map* tingkat kesukaran

Sebaran tingkat kesukaran butir soal pada setiap indikator kemampuan *Creative Problem Solving* peserta didik berdasarkan *Wright Map* dapat dilihat pada Gambar 6. Berdasarkan gambar tersebut, analisis difokuskan pada distribusi tingkat kesukaran item yang

ditunjukkan pada sisi kanan peta dalam skala logit. Hasil analisis menunjukkan bahwa tingkat kesukaran butir soal bervariasi dari kategori sangat mudah hingga sangat sukar. Butir soal yang tergolong sangat sukar adalah item 1b, 2b, dan 3b yang berada pada posisi logit tinggi, sehingga memerlukan kemampuan berpikir tingkat tinggi untuk menyelesaikannya. Sebagian butir soal berada pada kategori sukar, seperti item 2d, 2e, 2f, 3c, dan 3d yang berada pada rentang logit sedang ke atas. Sementara itu, mayoritas butir soal tergolong mudah, seperti item 1a, 1c, 1d, 1e, 1f, 2a, 2c, 3a, dan 3e yang berada pada rentang logit rendah hingga sedang. Adapun item 3f berada pada posisi logit paling rendah sehingga termasuk dalam kategori sangat mudah dan dapat diselesaikan oleh hampir seluruh peserta didik.

Variasi tingkat kesukaran butir soal yang tersebar dari kategori mudah hingga sulit menunjukkan bahwa instrumen yang digunakan memiliki kemampuan yang baik dalam mengukur berbagai tingkat kemampuan *Creative Problem Solving* peserta didik secara komprehensif.

Kesimpulan

Instrumen pengukuran kemampuan *Creative Problem Solving* (CPS) pada materi energi alternatif yang dikembangkan dalam penelitian ini terbukti memiliki kualitas yang baik berdasarkan analisis Model *Rasch*. Instrumen memenuhi asumsi unidimensionalitas dengan nilai *raw variance explained* sebesar 51,2% dan *unexplained variance in 1st contrast* sebesar 8,9%. Seluruh butir soal dinyatakan valid dan *fit* terhadap model, dengan tingkat reliabilitas yang tinggi yang ditunjukkan oleh nilai *person reliability* sebesar 0,84, *item reliability* sebesar 0,81, serta *Cronbach's Alpha* sebesar 0,93. Temuan ini menunjukkan bahwa instrumen mampu menghasilkan pengukuran yang konsisten, stabil, dan objektif dalam menilai kemampuan CPS peserta didik.

Sebaran tingkat kesukaran butir soal yang bervariasi dari kategori sangat mudah hingga sangat sukar menunjukkan bahwa instrumen mampu mengakomodasi berbagai tingkat kemampuan peserta didik secara komprehensif. Dengan demikian, instrumen ini layak digunakan sebagai alat evaluasi dalam pembelajaran fisika, khususnya pada materi energi alternatif. Pengembangan instrumen serupa pada materi lain serta dengan jumlah sampel yang lebih luas direkomendasikan untuk meningkatkan generalisasi hasil penelitian.

Ucapan Terimakasih

Penulis menyampaikan ucapan terima kasih kepada Universitas Pendidikan Indonesia, khususnya Program Studi Pendidikan Fisika, atas dukungan akademik yang diberikan selama proses penyusunan artikel ini. Penulis juga mengucapkan terima kasih kepada kepala salah satu SMA

swasta di Bandung yang telah memberikan izin sebagai lokasi penelitian. Selain itu, apresiasi turut disampaikan kepada seluruh pihak yang telah berkontribusi dalam pelaksanaan penelitian ini.

Referensi

- Dinantika, H. K., Suyanto, E., & Nyeneng, I. D. P. (2019). Pengaruh Penerapan Model Pembelajaran Project Based Learning Terhadap Kreativitas Siswa Pada Materi Energi Terbarukan. *Titian Ilmu: Jurnal Ilmiah Multi Sciences*, 11(2), 73–80. <https://doi.org/10.30599/jti.v11i2.473>
- Hu, R., Xiaohui, S., & Shieh, C.-J. (2017). A Study on the Application of Creative Problem Solving Teaching to Statistics Teaching. *EURASIA Journal of Mathematics, Science and Technology Education*, 13(7). <https://doi.org/10.12973/eurasia.2017.00708a>
- Junita, S. (2016). PENINGKATAN KEMAMPUAN CREATIVE PROBLEM SOLVING MATEMATIS SISWA SMP DENGAN PENDEKATAN CHALLENGE BASED LEARNING. *Jurnal Pengajaran Matematika Dan Ilmu Pengetahuan Alam*, 21(1), 19–23. <https://doi.org/10.18269/jpmipa.v21i1.36250>
- Khalid, M., Saad, S., Abdul Hamid, S. R., Ridhuan Abdullah, M., Ibrahim, H., & Shahrill, M. (2020). ENHANCING CREATIVITY AND PROBLEM SOLVING SKILLS THROUGH CREATIVE PROBLEM SOLVING IN TEACHING MATHEMATICS. *Creativity Studies*, 13(2), 270–291. <https://doi.org/10.3846/cs.2020.11027>
- Khotimah, K., Hadi, T. S., & Lestari, I. (2024). Application of The Rasch Model in Research Publications: A Bibliometric Analysis. *Plusminus: Jurnal Pendidikan Matematika*, 4(2), 229–240. <https://doi.org/10.31980/plusminus.v4i2.1466>
- Larson, L. C., & Miller, T. N. (2011). 21st Century Skills: Prepare Students for the Future. *Kappa Delta Pi Record*, 47(3), 121–123. <https://doi.org/10.1080/00228958.2011.10516575>
- Marjiastuti, K., & Wahyuni, S. (2015). ANALISIS KEMAMPUAN PESERTA DIDIK DENGAN MODEL RASCH. Retrieved from <https://api.semanticscholar.org/CorpusID:148672799>
- Neva Lionitha Ibrahim. (2025). Analisis Literatur Penggunaan Model Rasch untuk Validasi Instrumen Asesmen Pendidikan: Tren, Temuan, dan Implikasi. *Lencana: Jurnal Inovasi Ilmu Pendidikan*, 3(2), 46–60. <https://doi.org/10.55606/lencana.v3i2.5012>
- Nuraziza, R., & Suwama, I. R. (2018). MENGGALI KETERAMPILAN CREATIVE PROBLEM SOLVING YANG DIMILIKI SISWA SMP MELALUI PEMBELAJARAN IPA BERBASIS STEM. *WaPfi (Wahana Pendidikan Fisika)*, 3(1), 55. <https://doi.org/10.17509/wapfi.v3i1.10941>
- Perdana, S. A. (2018). ANALISIS KUALITAS INSTRUMEN PENGUKURAN PEMAHAMAN KONSEP PERSAMAAN KUADRAT MELALUI TEORI TES KLASIK DAN RASCH MODEL. *Jurnal Kiprah*, 6(1), 41–48. <https://doi.org/10.31629/kiprah.v6i1.574>
- Permatasari, N. (2025). Analisis kualitas instrumen penilaian materi keanekaragaman hayati melalui tes klasik dan Rasch model. *Bio-Pedagogi*, 14(1), 10. <https://doi.org/10.20961/bio-pedagogi.v14i1.88470>
- Prieto, G. A., & Delgado, A. R. (2003). Análisis de un test mediante el modelo de Rasch. *Psicothema*, 15, 94–100.
- Sumintono, B., & Widhiarso, W. (2015). Aplikasi pemodelan Rasch pada assessment pendidikan. Retrieved from <https://api.semanticscholar.org/CorpusID:147575669>
- Treffinger, D. J., & Isaksen, S. G. (2005). Creative Problem Solving: The History, Development, and Implications for Gifted Education and Talent Development. *Gifted Child Quarterly*, 49(4), 342–353. <https://doi.org/10.1177/001698620504900407>