



Analisis Instrumen Penilaian Keterampilan Proses Sains dan Creative Problem Solving Berbasis Model Rasch pada Materi Teori Kinetik Gas

Mira Fitri Yanti^{*}, Undang Rosidin, dan I Wayan Distrik
Magister Pendidikan Fisika, Universitas Lampung, Lampung, Indonesia
^{*} Email: fitrimira970@gmail.com

Abstrak

Keterampilan abad ke-21 saat ini menjadi perhatian utama dalam dunia pendidikan di tingkat global. Sehingga, penilaian keterampilan proses sains (KPS) dan creative problem solving (CPS) penting dilakukan untuk mengidentifikasi kemampuan peserta didik dalam berpikir ilmiah serta memecahkan masalah secara kreatif dalam pembelajaran fisika. Namun, keterbatasan instrumen yang dikembangkan oleh guru menyebabkan kedua keterampilan tersebut belum terukur secara optimal. Penelitian ini bertujuan untuk mengembangkan instrumen penilaian KPS dan CPS yang valid dan reliabel. Penelitian ini merupakan penelitian pengembangan (Research and Development) dengan menggunakan model ADDIE yang meliputi tahap analyze, design, develop, implementation, dan evaluation. Instrumen yang dikembangkan berupa tes berbentuk soal pilihan ganda untuk mengukur KPS dan CPS. Uji validitas dilakukan oleh tiga orang ahli pada bidangnya berdasarkan aspek materi, konstruk, dan bahasa, sedangkan uji empiris melibatkan 66 peserta didik dan dianalisis menggunakan perangkat lunak Winstep dengan pendekatan model Rasch. Hasil validasi ahli menunjukkan rata-rata sebesar 88,66% untuk KPS dan 86,66% untuk CPS dengan kategori sangat tinggi. Hasil uji coba menghasilkan 15 butir soal valid untuk KPS dan 14 butir soal valid untuk CPS, dengan reliabilitas sangat baik yang ditunjukkan oleh nilai Cronbach's alpha pada instrumen KPS sebesar 0,95, sedangkan pada CPS sebesar 0,96. Dengan demikian, kesimpulan pada penelitian ini ialah instrumen yang dikembangkan telah memenuhi kriteria kelayakan secara teoritis dan empiris, sehingga dapat digunakan sebagai alat penilaian untuk mengukur kemampuan KPS dan CPS peserta didik secara akurat.

Kata kunci: Instrumen Penilaian, Kemampuan Proses Sains, *Creative Problem Solving*, Teori Kinetik Gas

Analysis of Assessment Instrument for Science Process Skills and Creative Problem Solving Based on the Rasch Model in the Topic of Kinetic Theory of Gases

Abstract

21st-century skills have become a major focus in global education. Therefore, the assessment of science process skills (SPS) and creative problem solving (CPS) is essential to identify students' abilities in scientific thinking and creative problem-solving in physics learning. However, the limited availability of assessment instruments developed by teachers has resulted in these skills not being optimally measured. This study aims to develop a valid and reliable assessment instrument for SPS and CPS. This research employed a Research and Development (R&D) method using the ADDIE model, which consists of the stages of analysis, design, development, implementation, and evaluation. The instrument developed was in the form of multiple-choice test items to measure SPS and CPS. Content validity was assessed by three experts based on aspects of material, construction, and language, while empirical testing involved 66 students and was analyzed using Winsteps software with the Rasch model approach. The results of expert validation showed average scores of 88.66% for SPS and 86.66% for CPS, both categorized as very high. The trial results produced 15 valid items for SPS and 14 valid items for CPS, with very high reliability indicated by Cronbach's alpha values of 0.95 for SPS and 0.96 for CPS. In conclusion, the developed instrument

has met the criteria of theoretical and empirical validity, and therefore can be used as an assessment tool to accurately measure students' SPS and CPS abilities.

Keywords: *Assessment Instrument, Scientific Process Skills, Creative Problem Solving, Kinetic Theory of Gases*

Histori Naskah

Diserahkan: 25 Juni 2025

Direvisi: 19 April 2026

Diterima: 26 April 2026

How to cite:

Yanti, M.F., Rosidin, U., dan Distrik, I.W. (2026). Analisis Instrumen Penilaian Keterampilan Proses Sains dan Creative Problem Solving Berbasis Model Rasch pada Materi Teori Kinetik Gas. *Jurnal Ilmu Pendidikan dan Pembelajaran*, 4(2), 252-266. DOI: <https://doi.org/10.58706/jipp.v4n2.p252-266>.

PENDAHULUAN

Keterampilan abad ke-21 saat ini menjadi perhatian utama dalam dunia pendidikan di tingkat global. Peserta didik tidak hanya dituntut untuk menguasai pengetahuan konseptual, tetapi juga perlu memiliki kemampuan berpikir tingkat tinggi yang memungkinkan peserta didik memahami fenomena secara mendalam, serta menyelesaikan berbagai permasalahan secara kreatif dan ilmiah (Anjiana dkk., 2025). Kemampuan berpikir tingkat tinggi menjadi komponen penting dalam pembelajaran karena berperan dalam melatih peserta didik untuk menganalisis informasi, mengevaluasi berbagai alternatif solusi, dan mengambil keputusan secara rasional ketika menghadapi permasalahan yang kompleks (Khoiriyah & Shaleh, 2025). Pengembangan dan pengukuran kemampuan berpikir tingkat tinggi perlu menjadi perhatian dalam proses pembelajaran agar peserta didik mampu menerapkan pengetahuan yang dimiliki secara efektif dalam berbagai situasi nyata.

Kemampuan berpikir tingkat tinggi dapat tercermin melalui keterampilan proses sains (KPS) dan kemampuan *creative problem solving* (CPS) (Divona & Mustika, 2025). Kedua keterampilan ini mendorong peserta didik untuk melakukan proses ilmiah sekaligus menghasilkan berbagai alternatif solusi terhadap permasalahan yang dihadapi. Keterampilan proses sains berperan dalam membantu peserta didik memahami fenomena melalui tahapan penyelidikan ilmiah, seperti mengamati, merumuskan masalah, merancang percobaan, serta menganalisis dan menginterpretasikan data (Ratnasari dkk., 2017; Wandyka dkk., 2024). Namun, penguasaan keterampilan proses sains peserta didik masih tergolong rendah. Penelitian yang dilakukan Setyaningrum & Ikhtiyari (2025), menunjukkan sebagian besar peserta didik memiliki keterampilan proses sains yang tergolong sangat rendah. Hal ini menunjukkan bahwa peserta didik masih memerlukan dukungan pembelajaran yang mampu melatih keterampilan ilmiah secara lebih sistematis.

Kemampuan *creative problem solving* juga memiliki peran penting dalam membantu peserta didik menghasilkan ide-ide yang beragam dan inovatif dalam menyelesaikan suatu permasalahan (Putri, 2025). Penelitian yang dilakukan oleh Marjusa dkk. (2025) menunjukkan bahwa penerapan model *creative problem solving* secara signifikan mampu meningkatkan kreativitas belajar peserta didik. Hasil penelitian tersebut menunjukkan bahwa pendekatan *creative problem solving* efektif dalam mendorong peserta didik untuk berpikir lebih kreatif dalam proses pembelajaran. Pendekatan ini mendorong peserta didik untuk mengeksplorasi berbagai kemungkinan solusi serta mengembangkan strategi pemecahan masalah secara kreatif (Susanti, 2026). Namun demikian, dalam praktik pembelajaran konvensional, keterampilan ini sering kali terabaikan, sehingga peserta didik belum memperoleh ruang optimal untuk mengembangkan kemampuan berpikir ilmiahnya. Oleh karena itu, guru perlu mengintegrasikan dan mengoptimalkan keterampilan proses sains dalam pembelajaran IPA secara lebih terstruktur dan sistematis (Hasan dkk., 2024).

Keterampilan proses sains dan *creative problem solving* pada dasarnya saling berkaitan dalam mendukung pengembangan kemampuan berpikir tingkat tinggi dalam pembelajaran sains. Keterkaitan tersebut dapat ditinjau dari indikator yang digunakan dalam kedua keterampilan tersebut. Indikator keterampilan proses sains yang digunakan meliputi kemampuan menginterpretasi, melakukan prediksi, melakukan percobaan, observasi, klasifikasi, merumuskan hipotesis, serta mengajukan pertanyaan (Sund & Trowbridge, 1973). Sementara itu, indikator *creative problem solving* yang digunakan merujuk pada pendapat Treffinger & Isaksen, yaitu *fact finding*, *fact interpreting*, *idea finding*, *idea developing*, *solution generating*, dan *solution evaluating* (Treffinger & Isaksen, 2005). Secara konseptual, kedua keterampilan tersebut memiliki kesamaan dalam hal

proses berpikir yang menekankan pada identifikasi masalah, pengolahan informasi, serta pengembangan solusi secara sistematis.

Kesamaan karakteristik antara KPS dan CPS menunjukkan bahwa keduanya memerlukan pengukuran secara tepat dalam pembelajaran sains (Dinova & Mustika, 2025). Namun, penilaian terhadap KPS dan CPS masih menjadi tantangan dalam praktik pembelajaran. Penelitian yang dilakukan Hanan dkk. (2025) menunjukkan bahwa meskipun implementasi Kurikulum Merdeka telah mengarahkan pembelajaran pada pengembangan keterampilan dan karakter peserta didik, praktik penilaian yang dilakukan masih cenderung berfokus pada aspek kognitif, khususnya pada pemahaman dan analisis konsep. Hal ini menyebabkan KPS dan CPS kurang terukur secara optimal, meskipun kedua keterampilan tersebut telah dilatihkan dalam proses pembelajaran (Riatma & Susanti, 2025; Sukma dkk., 2026). Pengukuran keterampilan tersebut memerlukan instrumen penilaian yang mampu mengukur KPS dan CPS secara valid dan reliabel dengan pendekatan analisis yang mampu mengevaluasi kualitas butir soal dan kemampuan peserta didik secara objektif.

Analisis instrumen dengan pendekatan *Rasch Model* dipilih karena mampu mengevaluasi validitas dan reliabilitas instrumen penelitian secara komprehensif, sehingga kualitas butir soal dan kemampuan peserta didik dapat dianalisis secara objektif. Model ini telah digunakan secara luas dalam berbagai penelitian untuk menganalisis kualitas butir soal dan mengidentifikasi karakteristik peserta didik (Ling dkk., 2014; Matore dkk., 2018; Zamri dkk., 2015). *Rasch Model* memungkinkan pengujian kelayakan instrumen, reliabilitas pengukuran, identifikasi bias item, serta pengujian unidimensionalitas suatu konstruk (Suryani dkk., 2015). Selain itu, model ini juga dapat mengestimasi kemampuan peserta didik secara objektif dengan asumsi bahwa probabilitas menjawab benar meningkat seiring dengan peningkatan kemampuan siswa (Islam dkk., 2020; Xiao dkk., 2018).

Penelitian terdahulu telah banyak mengembangkan instrumen penilaian untuk mengukur keterampilan proses sains maupun *creative problem solving* pada pembelajaran. Pengembangan instrumen KPS telah dilakukan pada beberapa materi, seperti optika, tekanan, getaran dan gelombang, tumbukan, dan gerak pada jenjang SMP (Nurhayati dkk., 2019; Suryani dkk., 2015). Sementara, instrumen untuk CPS telah dikembangkan untuk tingkat perguruan tinggi, namun tidak terfokus pada materi fisika. Sementara itu, pengembangan instrumen *creative problem solving* lebih banyak dilakukan pada tingkat perguruan tinggi maupun pada konteks matematika dan sains secara umum (Yurt, 2025; Park & Kang, 2012). Namun, kajian yang secara khusus mengembangkan instrumen untuk mengukur kemampuan *creative problem solving* pada materi fisika di jenjang SMA masih terbatas, terutama pada materi yang bersifat abstrak seperti teori kinetik gas. Penggunaan pendekatan analisis modern, seperti pemodelan Rasch, dalam pengembangan instrumen juga belum banyak dimanfaatkan.

Untuk mendukung landasan pengembangan instrumen, peneliti telah melakukan studi lapangan melalui wawancara dengan guru mata pelajaran fisika kelas XI di dua sekolah negeri di Kabupaten Tanggamus, Lampung. Hasil wawancara menunjukkan bahwa pembelajaran fisika, khususnya pada materi teori kinetik gas, telah melibatkan peserta didik secara aktif. Berbagai model pembelajaran seperti *Problem-Based Learning (PBL)*, *Project-Based Learning (PjBL)*, *STEM*, dan *inkuiri* telah diterapkan, termasuk dalam kegiatan percobaan sederhana. Hal ini mengindikasikan bahwa keterampilan proses sains dan kemampuan pemecahan masalah kreatif telah diintegrasikan dalam proses pembelajaran. Namun, para guru belum memiliki instrumen khusus yang valid dan reliabel untuk mengukur kedua keterampilan tersebut. Instrumen yang digunakan saat ini belum melewati proses validasi empiris, sehingga belum dapat dipastikan kelayakannya sebagai alat ukur.

METODE

Penelitian ini merupakan penelitian pengembangan (*Research and Development*) yang bertujuan menghasilkan instrumen penilaian keterampilan proses sains dan *creative problem solving*. Prosedur pengembangan dalam penelitian ini berpedoman pada model pengembangan instruksional ADDIE yang dikembangkan oleh Branch (2009) yang terdiri dari lima langkah, yaitu *Analyze*, *Design*, *Development*, *Implementation*, dan *Evaluation*. Prosedur penelitian dapat diilustrasikan pada Gambar 1.



Gambar 1. Langkah Pengembangan ADDIE
(Sumber: Adaptasi dari Branch, 2009)

Berdasarkan Gambar 1, langkah pengembangan dimulai dengan tahap *Analyze* yang dilakukan melalui studi pendahuluan dengan wawancara kepada guru fisika kelas XI di dua sekolah menengah negeri di Kabupaten Tanggamus untuk mengidentifikasi kebutuhan instrumen dan permasalahan penilaian. Tahap *Design* meliputi penyusunan kisi-kisi (*table of specification*) berdasarkan indikator keterampilan proses sains menurut Sund & Trowbridge (1973), yaitu menginterpretasi, prediksi, observasi, klasifikasi, hipotesis, melakukan percobaan, dan mengajukan pertanyaan, sedangkan indikator *creative problem solving* menurut Treffinger & Isaksen (2005), yaitu *fact finding*, *fact interpreting*, *idea finding*, *idea developing*, *solution generating*, dan *solution evaluating*. Berdasarkan indikator tersebut disusun blueprint instrumen yang memuat keterkaitan antara indikator, materi, dan level kognitif.

Pada tahap *Development*, dikembangkan 30 butir soal pilihan ganda yang terdiri dari 15 soal keterampilan proses sains dan 15 soal *creative problem solving*, kemudian divalidasi oleh tiga validator yang terdiri dari dua dosen pendidikan fisika dan satu guru fisika SMA dengan kriteria memiliki latar belakang bidang fisika serta pengalaman dalam pembelajaran atau pengembangan instrumen. Validasi dilakukan menggunakan angket skala *Likert* yang mencakup aspek konstruk, materi, dan bahasa. Tahap *Implementation* dilakukan melalui uji coba skala terbatas kepada 33 peserta didik dan uji coba skala luas kepada 66 peserta didik kelas XI dengan kemampuan yang heterogen. Penggunaan dua tahap uji coba bertujuan untuk memperoleh gambaran awal kualitas instrumen serta menguji konsistensi hasil pengukuran pada jumlah responden yang lebih besar. Tahap *Evaluation* dilakukan secara berkelanjutan berdasarkan hasil validasi ahli dan analisis empiris untuk menilai kelayakan instrumen.

Analisis data dilakukan menggunakan pendekatan *Rasch Model* dengan bantuan perangkat lunak *Winsteps* untuk mengevaluasi kualitas instrumen secara komprehensif. Analisis meliputi uji validitas konstruk melalui *item fit* dengan kriteria nilai *infit* dan *outfit MNSQ* berada pada rentang 0,5–1,5, reliabilitas instrumen melalui *person* dan *item reliability* dengan nilai $\geq 0,70$, serta uji unidimensionalitas melalui *principal component analysis* (PCA) dengan kriteria varians yang dijelaskan oleh model $\geq 20\%$. Selain itu, dilakukan analisis tingkat kesukaran, daya beda, serta deteksi bias butir (*differential item functioning*). Jumlah responden dalam penelitian ini telah memenuhi kriteria minimum dalam analisis *Rasch* untuk menghasilkan estimasi parameter yang stabil (Linacre, 1994).

HASIL DAN PEMBAHASAN

Instrumen yang dikembangkan dalam penelitian ini berjumlah 30 butir soal, yang terdiri atas 15 butir soal keterampilan proses sains (KPS) dan 15 butir soal *creative problem solving* (CPS). Instrumen disusun dalam bentuk pilihan ganda yang dirancang untuk mengukur kemampuan berpikir tingkat tinggi pada materi teori kinetik gas. Instrumen ini dikembangkan dengan mengacu pada model ADDIE, yaitu *analyze*, *design*, *development*, *implementation*, dan *evaluation*.

Tahap *Analyze*

Tahap *analyze* bertujuan untuk mengumpulkan dan menganalisis berbagai informasi yang relevan, serta merumuskan permasalahan yang akan menjadi landasan dalam pengembangan instrumen tes KPS dan CPS. Pada tahap ini dilakukan analisis kebutuhan dengan penyebaran angket dan wawancara kepada para guru. Responden dalam kegiatan ini adalah guru-guru yang berasal dari SMA Negeri di Kabupaten Tanggamus,

Lampung. Hasil analisis kebutuhan menunjukkan bahwa guru belum memiliki instrumen khusus untuk mengukur KPS dan CPS secara terintegrasi, serta belum memanfaatkan analisis Rasch dalam evaluasi instrumen.

Temuan dalam analisis kebutuhan menunjukkan adanya ketidaksesuaian antara praktik pembelajaran yang telah mengarah pada pengembangan keterampilan berpikir tingkat tinggi dengan sistem penilaian yang digunakan di kelas. Hal ini sejalan dengan penelitian yang dilakukan oleh Dwikoranto dkk. (2025) yang menunjukkan bahwa keterampilan berpikir tingkat tinggi peserta didik masih tergolong rendah yang disebabkan oleh belum optimalnya instrumen penilaian yang digunakan untuk mengukur keterampilan tersebut. Guru cenderung masih menggunakan instrumen yang berfokus pada penguasaan konsep, sehingga proses ilmiah dan kemampuan pemecahan masalah kreatif peserta didik belum terukur secara optimal (Dwikorato dkk., 2025). Hal ini membuktikan urgensi pengembangan instrumen yang kemudian dijadikan dasar untuk melanjutkan ke tahap berikutnya, yaitu tahap *Design*.

Tahap Design

Tahap *design* peneliti dilakukan untuk perancangan desain instrumen yang akan dikembangkan. Pada tahap ini instrumen yang dirancang akan melalui proses uji validitas, reliabilitas, tingkat kesukaran, serta daya beda. Setiap butir dan format instrumen dikembangkan dengan mencantumkan Kompetensi Dasar (KD), indikator soal, level kognitif, serta aspek-aspek yang merepresentasikan KPS dan CPS. Indikator *creative problem solving* yang dikembangkan pada penelitian ini merujuk pada pendapat Treffinger & Isaksen (2005) dan indikator keterampilan proses sains mengacu pada Sund & Trowbridge (1973), yang ditunjukkan pada Tabel 1.

Tabel 1. Indikator KPS dan CPS

No.	Indikator	Keterangan
Keterampilan Proses Sains (KPS)		
1.	Observasi	Peserta didik mampu mengamati objek atau fenomena secara sistematis dengan menggunakan indera atau alat bantu untuk memperoleh data yang akurat.
2.	Mengajukan pertanyaan	Peserta didik mampu menyusun pertanyaan yang relevan dan bermakna untuk menggali informasi lebih lanjut terkait suatu fenomena atau masalah.
3.	Klasifikasi	Peserta didik mampu mengelompokkan objek, data, atau fenomena berdasarkan kesamaan dan perbedaan karakteristik tertentu.
4.	Menginterpretasi	Peserta didik mampu menafsirkan data atau informasi yang diperoleh.
5.	Prediksi	Peserta didik mampu memperkirakan kejadian atau hasil yang akan terjadi.
6.	Hipotesis	Peserta didik mampu merumuskan dugaan sementara yang logis dan dapat diuji berdasarkan hasil pengamatan atau pengetahuan awal.
7.	Melakukan percobaan	Peserta didik mampu merancang dan melaksanakan kegiatan eksperimen untuk menguji suatu hipotesis yang telah dirumuskan.
Creative Problem Solving (CPS)		
1.	<i>Fact finding</i> (penemuan fakta)	Peserta didik mampu mengumpulkan informasi tentang situasi yang bermasalah. Mengeksplorasi dan mengidentifikasi fakta-fakta tersebut.
2.	<i>Fact interpreting</i> (interpretasi fakta)	Peserta didik mampu mengolah dan menafsirkan fakta yang telah diperoleh untuk memahami inti permasalahan.
3.	<i>Idea finding</i> (pencarian ide)	Peserta didik mampu menghasilkan berbagai alternatif ide atau solusi yang dapat digunakan untuk menyelesaikan masalah.
4.	<i>Idea developing</i> (pengembangan ide)	Peserta didik mampu mengembangkan ide-ide yang telah dihasilkan agar menjadi solusi yang lebih jelas dan aplikatif.
5.	<i>Solution Generating</i> (pemilihan solusi)	Peserta didik mampu menentukan solusi terbaik dari berbagai ide alternatif yang telah dikembangkan untuk menyelesaikan masalah.
6.	<i>Solution Evaluating</i> (evaluasi solusi)	Peserta didik mampu menilai efektivitas solusi yang dipilih, serta mempertimbangkan kelebihan dan kekurangannya dalam menyelesaikan masalah.

(Sumber: Sund & Trowbridge, 1973; Treffinger & Isaksen, 2005)

Berdasarkan indikator pada Tabel 1, menunjukkan bahwa KPS dan CPS pada dasarnya berada dalam satu kerangka kemampuan berpikir tingkat tinggi yang saling melengkapi dalam pembelajaran sains. KPS berperan sebagai dasar dalam membangun pemahaman ilmiah melalui kegiatan penyelidikan, sedangkan CPS memperluas proses tersebut ke arah pengembangan ide dan pengambilan keputusan yang kreatif. Oleh karena

itu, pengintegrasian kedua keterampilan ini dalam penilaian menjadi penting agar kemampuan peserta didik dapat diukur secara lebih komprehensif, tidak hanya dari aspek pemahaman konsep, tetapi juga dari kemampuan berpikir ilmiah dan kreatif dalam menyelesaikan masalah.

Tahap Develop

Tahap *Develop* dilakukan pengembangan instrumen berupa 30 butir soal, yang terdiri dari 15 butir soal KPS dan 15 butir soal CPS. Instrumen yang telah disusun kemudian divalidasi oleh tiga validator yang terdiri atas dua dosen pendidikan fisika dan satu guru fisika SMA. Validator dipilih berdasarkan kriteria memiliki latar belakang pendidikan di bidang fisika atau pendidikan fisika, serta memiliki pengalaman dalam pengembangan instrumen atau pembelajaran fisika. Validasi dilakukan menggunakan lembar validasi berbentuk skala *Likert* 4 poin yang mencakup aspek konstruk, materi, dan bahasa. Hasil validasi dianalisis untuk menentukan tingkat kelayakan instrumen sebelum dilakukan uji coba kepada peserta didik, yang disajikan pada Tabel 2.

Tabel 2. Hasil Uji Validasi Ahli

No	Aspek	Persentase skor rata-rata instrumen		Kategori
		KPS	CPS	
1.	Materi	88,86	86,66	Sangat Tinggi
2.	Konstruksi	87,33	86,88	Sangat Tinggi
3.	Bahasa	89,77	86,33	Sangat Tinggi
Rata-rata Keseluruhan		88,66	86,62	Sangat Tinggi

Berdasarkan Tabel 2, hasil uji validasi ahli menunjukkan bahwa instrumen yang dikembangkan memiliki kualitas yang sangat baik pada seluruh aspek penilaian, baik pada keterampilan proses sains (KPS) maupun *creative problem solving* (CPS). Aspek materi memperoleh persentase sebesar 88,86 (KPS) dan 86,66 (CPS), yang mengindikasikan bahwa butir soal telah sesuai dengan konsep, indikator, serta materi teori kinetik gas yang diukur. Aspek konstruksi juga menunjukkan nilai tinggi, yaitu 87,33 (KPS) dan 86,88 (CPS), yang berarti penyusunan soal telah memenuhi kaidah penulisan soal yang baik, seperti kejelasan stem, kesesuaian opsi jawaban, serta tidak menimbulkan ambiguitas. Sementara itu, aspek bahasa memperoleh nilai tertinggi, yaitu 89,77 (KPS) dan 86,33 (CPS), yang menunjukkan bahwa penggunaan bahasa dalam instrumen telah komunikatif, sesuai dengan tingkat perkembangan peserta didik, serta tidak menimbulkan penafsiran ganda.

Hasil uji validasi ahli memiliki rata-rata skor validasi mencapai 88,66 untuk KPS dan 86,62 untuk CPS dengan kategori “sangat tinggi”, sehingga instrumen dinyatakan layak untuk digunakan dalam tahap uji coba. Hasil ini menunjukkan bahwa instrumen telah memenuhi validitas isi dan konstruk, yaitu kesesuaian antara butir soal dengan indikator yang diukur (Arikunto, 2011). Validasi oleh ahli merupakan tahap penting dalam pengembangan instrumen untuk memastikan bahwa instrumen benar-benar mengukur kemampuan yang seharusnya diukur serta memiliki kualitas yang baik sebelum diujicobakan secara empiris (Ramadhan dkk., 2024). Dengan demikian, tingginya hasil validasi ini menunjukkan bahwa instrumen yang dikembangkan telah memiliki dasar teoretis dan teknis yang kuat untuk digunakan dalam mengukur keterampilan proses sains dan *creative problem solving* peserta didik.

Tahap Implementation

Instrumen yang telah dinyatakan layak oleh validator kemudian diuji cobakan kepada peserta didik pada dua SMA Negeri di Kabupaten Tanggamus, Lampung. Uji coba dilakukan dalam dua tahap, yaitu uji coba skala terbatas dan uji coba skala luas yang melibatkan sebanyak 66 peserta didik. Responden merupakan peserta didik kelas XI dengan karakteristik heterogen dari segi kemampuan akademik. Penggunaan dua tahap uji coba bertujuan untuk memperoleh gambaran awal kualitas instrumen pada skala kecil, serta menguji konsistensi hasil pengukuran pada jumlah responden yang lebih besar sesuai dengan prinsip dalam analisis Rasch yang memerlukan ukuran sampel yang memadai untuk menghasilkan estimasi parameter yang stabil.

Hasil Uji Coba Terbatas

Hasil analisis Rasch uji coba terbatas merujuk pada tiga parameter yaitu: Nilai *Outfit Mean Square* (MNSQ) yang diterima : $0,5 < MNSQ < 1,5$. Nilai *Outfit Z-Standard* (ZSTD) yang diterima: $-2,0 < ZSTD < +2,0$, dan Nilai *Pt Mean Corr* yang diterima : $0,4 < Pt Measure Corr < 0,85$. (Boone, Staver, & Yale; 2014). Nilai MNSQ menunjukkan kesesuaian data dengan model. Semakin mendekati nilai 1,00 maka semakin sesuai

butir soal tersebut. Nilai ZSTD bertindak sebagai uji-t yang menunjukkan hipotesis kesesuaian data. Nilai ZSTD yang diharapkan adalah mendekati (Sumintono & Widhiarso, 2015). *Pt Mean Corr* mengacu pada korelasi antara skor peserta didik dan kemampuan, semakin tinggi dan semakin positif korelasinya maka semakin baik (Wei dkk., 2012). Hasil uji coba skala terbatas pada instrumen CPS ditunjukkan pada Gambar 2.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	ITEM
10	21	33	-.71	.40	1.66	3.42	1.90	3.25	A-.26	.41	39.4	70.5	Q10
2	13	33	.48	.39	1.03	.24	1.16	.79	B .35	.40	69.7	68.6	Q2
7	18	33	-.25	.38	1.06	.46	1.03	.25	C .36	.42	66.7	69.3	Q7
4	19	33	-.40	.39	1.05	.39	1.00	.04	D .38	.42	69.7	69.5	Q4
15	20	33	-.55	.39	.99	.01	1.04	.24	E .41	.41	72.7	70.1	Q15
13	15	33	.19	.38	1.02	.17	1.00	.08	F .40	.41	66.7	68.6	Q13
8	15	33	.19	.38	1.01	.13	1.01	.10	G .40	.41	72.7	68.6	Q8
14	21	33	-.71	.40	1.00	.06	1.00	.09	H .40	.41	69.7	70.5	Q14
6	14	33	.33	.39	.98	-.08	.97	-.12	g .43	.41	75.8	68.4	Q6
11	11	33	.79	.40	.96	-.16	.92	-.25	f .43	.39	69.7	71.0	Q11
12	15	33	.19	.38	.89	-.70	.84	-.84	e .53	.41	72.7	68.6	Q12
1	16	33	.04	.38	.88	-.76	.88	-.67	d .53	.42	78.8	68.8	Q1
3	16	33	.04	.38	.86	-.90	.81	-1.08	c .56	.42	72.7	68.8	Q3
5	16	33	.04	.38	.80	-1.36	.75	-1.50	b .61	.42	72.7	68.8	Q5
9	14	33	.33	.39	.79	-1.48	.73	-1.50	a .62	.41	81.8	68.4	Q9
MEAN	16.3	33.0	.00	.39	1.00	-.04	1.00	-.07			70.1	69.2	
P.SD	2.9	.0	.43	.01	.20	1.10	.27	1.10			9.1	.8	

Gambar 2. Analisis Rasch Hasil Uji Coba Terbatas Instrumen CPS

Berdasarkan hasil analisis pada Gambar 2, diperoleh sebanyak 14 butir soal dinyatakan valid dan 1 butir soal dinyatakan tidak valid, yaitu soal nomor 10. Ketidakvalidan butir soal tersebut disebabkan karena tidak memenuhi ketiga parameter yang telah ditetapkan. Dalam penelitian ini, suatu butir soal dinyatakan valid apabila memenuhi minimal dua dari tiga parameter yang digunakan. Selain itu, analisis daya beda yang ditinjau melalui nilai *Model Standard Error* (SE) menunjukkan bahwa seluruh butir soal berada pada kategori baik, karena memiliki nilai SE kurang dari 0,5. Hal ini mengindikasikan bahwa setiap butir soal memiliki kemampuan yang baik dalam membedakan tingkat kemampuan peserta didik (Purniasari dkk., 2021). Hal ini tidak terlepas pada tingkat kesukaran butir soal yang disajikan pada Tabel 3.

Tabel 3. Hasil Uji Tingkat Kesukaran Butir Soal CPS

Kriteria	Kategori	Butir Soal
> 0,43	Sangat Sukar	S2, S11
0 - 0,43	Sukar	S1, S3, S5, S6, S8, S9, S12, S13
- 0,43 - 0	Mudah	S7, S4
< - 0,43	Sangat Mudah	S10, S15, S14

Berdasarkan Tabel 3, tingkat kesukaran butir soal dianalisis menggunakan nilai *measure logit* dan simpangan baku (SD) item. Soal yang baik adalah soal yang tidak terlalu mudah maupun terlalu sulit (Arikunto, 2011). Berdasarkan hasil analisis pada Gambar 2 diperoleh nilai standar deviasi (P.SD) sebesar 0,43, yang kemudian digunakan sebagai dasar pengelompokan tingkat kesukaran menjadi empat kategori, yaitu sangat sukar, sukar, mudah, dan sangat mudah (Sumintono & Widhiarso, 2015). Hasil ini menunjukkan bahwa butir soal yang dikembangkan memiliki variasi tingkat kesukaran yang baik, sehingga mampu mengakomodasi perbedaan kemampuan peserta didik secara lebih proporsional. Hal ini juga didukung oleh hasil analisis reliabilitas yang ditunjukkan pada Gambar 3.

SUMMARY OF 33 MEASURED PERSON								
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	7.6	15.0	.05	.59	1.00	.08	.99	.07
SEM	.6	.0	.18	.01	.02	.11	.03	.12
P.SD	3.2	.0	1.01	.06	.12	.60	.18	.68
S.SD	3.2	.0	1.02	.06	.12	.61	.18	.69
MAX.	13.0	15.0	1.95	.77	1.33	1.58	1.48	1.82
MIN.	2.0	15.0	-1.95	.53	.80	-1.43	.66	-1.43
REAL RMSE	.61	TRUE SD	.81	SEPARATION	1.33	PERSON RELIABILITY	.64	
MODEL RMSE	.59	TRUE SD	.81	SEPARATION	1.37	PERSON RELIABILITY	.65	
S.E. OF PERSON MEAN = .18								
PERSON RAW SCORE-TO-MEASURE CORRELATION = 1.00								
CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .69 SEM = 1.78								
STANDARDIZED (50 ITEM) RELIABILITY = .86								

Gambar 3. Analisis Rasch Hasil Reliabilitas Uji Coba Terbatas Instrumen CPS

Berdasarkan Gambar 3, nilai *Cronbach's alpha* sebesar 0,69 menunjukkan bahwa instrumen CPS berada pada kategori cukup, sehingga hasil pengukuran dapat dianggap konsisten dan dapat diandalkan (Sumintono & Widhiarso, 2015). Selain itu, nilai separasi menggambarkan kemampuan instrumen dalam mengelompokkan responden berdasarkan tingkat kemampuannya. Semakin tinggi nilai separasi, maka semakin baik instrumen dalam membedakan kemampuan peserta didik (Sumintono & Widhiarso, 2015). Nilai separasi tersebut dapat dihitung menggunakan persamaan (1).

$$Strata(H) = \frac{[(4 \times separation) + 1]}{3} \tag{1}$$

Perhitungan menggunakan persamaan (1) menghasilkan nilai strata sebesar 2,1 yang kemudian dibulatkan menjadi 2. Hasil ini menunjukkan bahwa instrumen mampu mengelompokkan butir soal ke dalam dua kategori tingkat kesukaran, yaitu mudah dan sulit. Pengelompokan tersebut mengindikasikan bahwa instrumen telah memiliki kemampuan yang cukup dalam membedakan tingkat kesukaran soal, sehingga dapat digunakan untuk mengidentifikasi variasi kemampuan peserta didik secara lebih sederhana namun tetap informatif. Hasil uji coba skala terbatas terhadap instrumen keterampilan proses sains menunjukkan data sebagaimana disajikan pada Gambar 4.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	ITEM	
8	14	33	.49	.39	1.04	.33	1.24	1.10	A	.37	.44	75.8	69.0	Q23
13	16	33	.19	.39	1.03	.27	1.20	1.01	B	.38	.44	72.7	69.1	Q28
15	20	33	-.42	.40	1.05	.33	1.13	.62	C	.38	.44	69.7	71.3	Q30
1	18	33	-.11	.39	1.05	.37	1.12	.62	D	.39	.44	72.7	70.0	Q16
4	19	33	-.27	.39	1.09	.57	1.09	.48	E	.37	.44	69.7	70.7	Q19
7	21	33	-.58	.40	1.07	.47	1.05	.26	F	.37	.43	66.7	72.0	Q22
10	24	33	-1.09	.43	1.04	.26	1.00	.12	G	.37	.40	72.7	75.4	Q25
2	14	33	.49	.39	.99	-.01	1.02	.17	H	.44	.44	75.8	69.0	Q17
6	15	33	.34	.39	1.02	.21	1.01	.09	g	.43	.44	69.7	68.8	Q21
11	13	33	.64	.40	1.01	.13	1.00	.07	f	.42	.43	72.7	70.1	Q26
9	17	33	.04	.39	.98	-.10	.99	.00	e	.46	.45	69.7	69.6	Q24
14	19	33	-.27	.39	.92	-.48	.83	-.74	d	.53	.44	75.8	70.7	Q29
3	16	33	.19	.39	.91	-.52	.85	-.71	c	.53	.44	72.7	69.1	Q18
5	16	33	.19	.39	.89	-.71	.89	-.50	b	.54	.44	72.7	69.1	Q20
12	16	33	.19	.39	.81	-1.28	.77	-1.16	a	.61	.44	72.7	69.1	Q27
MEAN	17.2	33.0	.00	.39	.99	-.01	1.01	.10				72.1	70.2	
P.SD	2.9	.0	.45	.01	.08	.50	.13	.63				2.5	1.7	

Gambar 4. Analisis Rasch Hasil Uji Coba Terbatas Instrumen KPS

Berdasarkan hasil analisis menggunakan Rasch Model pada Gambar 4 diperoleh jumlah soal valid sebanyak 15 butir soal. Nilai SE pada setiap butir soal kurang dari 0,5. Hal ini menunjukan bahwa seluruh butir soal memiliki daya beda pada kategori tinggi. Hasil uji tingkat kesukaran pada instrumen KPS dipengaruhi

oleh nilai P.SD. Berdasarkan Gambar 4 di atas dapat dilihat nilai P.SD sebesar 0,45, sehingga tingkat kesukaran butir soal diperoleh pada Tabel 4.

Tabel 4. Hasil Uji Tingkat Kesukaran Butir Soal KPS

Kriteria	Kategori	Butir Soal
> 0,45	Sangat Sukar	S23, S17, S26
0 - 0,45	Sukar	S18, S20, S21, S24, S27, S28
- 0,45 - 0	Mudah	S16, S19, S29, S30
< - 0,45	Sangat Mudah	S22, S25

Berdasarkan Tabel 4, tingkat kesukaran butir soal KPS menunjukkan distribusi yang bervariasi pada empat kategori, yaitu sangat sukar, sukar, mudah, dan sangat mudah. Sebagian besar butir soal berada pada kategori sukar, diikuti oleh kategori mudah, sedangkan hanya beberapa butir yang termasuk dalam kategori sangat sukar dan sangat mudah. Distribusi ini menunjukkan bahwa instrumen memiliki variasi tingkat kesukaran yang cukup baik, sehingga mampu mengakomodasi perbedaan kemampuan peserta didik secara lebih proporsional. Adanya variasi ini penting untuk memastikan bahwa instrumen tidak hanya mengukur peserta didik dengan kemampuan tertentu, tetapi juga mampu membedakan kemampuan pada berbagai tingkat. Dengan demikian, kualitas instrumen dalam aspek tingkat kesukaran dapat dikatakan telah memenuhi kriteria soal yang baik, yaitu tidak terlalu mudah dan tidak terlalu sulit (Arikunto, 2011). Selanjutnya, uji reliabilitas instrumen KPS menggunakan model Rasch dapat dilihat pada Gambar 5 di bawah ini.

SUMMARY OF 33 MEASURED PERSON

	TOTAL	COUNT	MEASURE	MODEL	INFIT		OUTFIT	
	SCORE			S.E.	MNSQ	ZSTD	MNSQ	ZSTD
MEAN	7.8	15.0	.11	.60	1.00	.00	1.01	.06
SEM	.6	.0	.19	.01	.02	.11	.03	.12
P.SD	3.3	.0	1.05	.08	.11	.62	.18	.68
S.SD	3.3	.0	1.07	.08	.11	.63	.18	.69
MAX.	13.0	15.0	1.94	.77	1.25	2.07	1.46	2.07
MIN.	2.0	15.0	-1.94	.53	.82	-1.75	.70	-1.66
REAL RMSE	.61	TRUE SD	.85	SEPARATION	1.38	PERSON RELIABILITY	.66	
MODEL RMSE	.60	TRUE SD	.86	SEPARATION	1.43	PERSON RELIABILITY	.67	
S.E. OF PERSON MEAN = .19								

PERSON RAW SCORE-TO-MEASURE CORRELATION = 1.00
 CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .70 SEM = 1.77
 STANDARDIZED (50 ITEM) RELIABILITY = .87

Gambar 5. Analisis Rasch Hasil Reliabilitas Uji Coba Terbatas Instrumen KPS

Berdasarkan Gambar 5, nilai *Cronbach's alpha* sebesar 0,70 menunjukkan bahwa instrumen KPS memiliki reliabilitas pada kategori cukup, sehingga hasil pengukuran dapat dianggap konsisten dan dapat diandalkan (Sumintono & Widhiarso, 2015). Nilai separasi yang diperoleh sebesar 1,38 menunjukkan bahwa instrumen mampu membedakan tingkat kemampuan peserta didik dengan cukup baik. Hasil perhitungan nilai strata menggunakan persamaan (1) diperoleh sebesar 2,2 yang kemudian dibulatkan menjadi 2, sehingga butir soal dapat dikelompokkan ke dalam dua kategori tingkat kesukaran, yaitu mudah dan sulit. Hal ini menunjukkan bahwa instrumen memiliki kemampuan yang memadai dalam mengklasifikasikan tingkat kesukaran soal.

Hasil uji coba terbatas menunjukkan bahwa instrumen yang dikembangkan telah memenuhi kriteria kelayakan secara umum, baik ditinjau dari aspek validitas, reliabilitas, daya beda, maupun tingkat kesukaran. Pada instrumen *creative problem solving* (CPS), sebagian besar butir soal telah memenuhi parameter kesesuaian model Rasch, meskipun terdapat satu butir yang tidak valid karena tidak memenuhi kriteria MNSQ, ZSTD, dan *point measure correlation*. Sementara itu, seluruh butir pada instrumen keterampilan proses sains (KPS) dinyatakan valid dengan daya beda yang baik. Variasi tingkat kesukaran pada kedua instrumen juga menunjukkan distribusi yang proporsional, sehingga mampu mengakomodasi perbedaan kemampuan peserta didik. Nilai reliabilitas pada kedua instrumen berada pada kategori cukup, yang mengindikasikan bahwa hasil pengukuran relatif konsisten dan dapat diandalkan (Sumintono & Widhiarso, 2015). Temuan ini menunjukkan bahwa instrumen telah memiliki kualitas awal yang baik untuk digunakan dalam pengukuran, sehingga

sebanyak 14 butir soal *creative problem solving* dan 15 butir soal keterampilan proses sains yang dinyatakan valid kemudian dilanjutkan ke tahap uji coba skala luas guna memperoleh gambaran kualitas instrumen yang lebih komprehensif.

Hasil Uji Coba Luas

Uji coba skala luas menggunakan 66 responden yang terdiri dari dua sekolah. Hasil uji coba ini dimaksudkan untuk melihat kualitas dan kelayakan instrumen. Uji kelayakan meliputi uji validitas, reliabilitas, daya beda, dan tingkat kesukaran instrumen. Hasil yang diperoleh untuk instrumen keterampilan proses sains pada skala luas dapat dilihat pada Gambar 6.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	Item	
2	38	66	-1.25	.43	2.03	3.49	1.50	1.74	A	.62	.79	65.1	82.5	S2
9	32	66	-.23	.40	1.76	3.02	1.54	1.38	B	.64	.78	58.1	79.5	S9
7	33	66	-.39	.40	.99	.03	1.24	1.39	C	.77	.78	81.4	79.8	S7
12	25	66	.85	.39	1.41	1.95	1.17	.52	D	.67	.73	62.8	77.8	S12
6	25	66	.85	.39	1.29	1.43	1.01	.17	E	.69	.73	67.4	77.8	S6
1	29	66	.24	.39	.96	-.15	.71	-.84	F	.78	.76	72.1	78.6	S1
13	27	66	.54	.39	.96	-.14	.79	-.51	G	.76	.75	69.8	77.6	S13
5	29	66	.24	.39	.79	-1.05	.92	-.13	g	.79	.76	90.7	78.6	S5
8	38	66	-1.25	.43	.91	-.32	.65	-.58	f	.81	.79	83.7	82.5	S8
10	26	66	.70	.39	.71	-1.57	.54	-1.34	e	.80	.74	83.7	77.3	S10
14	31	66	-.07	.40	.65	-1.84	.62	-1.17	d	.83	.77	93.0	79.2	S14
3	33	66	-.39	.40	.60	-2.13	.44	-1.85	c	.85	.78	90.7	79.8	S3
11	32	66	-.23	.40	.57	-2.35	.45	-1.86	b	.85	.78	95.3	79.5	S11
4	28	66	.39	.39	.51	-2.98	.37	-1.94	a	.84	.75	90.7	78.1	S4
MEAN	30.4	66.0	.00	.40	1.01	-.19	1.00	-.26				78.9	79.2	
P.SD	4.1	.0	.66	.01	.44	1.94	.64	1.47				12.2	1.6	

Gambar 6. Analisis Rasch Hasil Uji Coba Luas Instrumen CPS

Berdasarkan Gambar 6 diperoleh data sebanyak 14 butir soal dinyatakan valid. 13 soal memenuhi ketiga parameter yang sudah ditetapkan, sedangkan soal nomor 9 hanya memenuhi 2 parameter namun masih termasuk pada kategori valid. Uji daya beda instrumen dapat dilihat pada kolom model S.E diperoleh seluruh nilai kurang dari 0,5 sehingga daya beda seluruh butir soal masuk pada kategori tinggi. Uji tingkat kesukaran instrumen beracuan pada nilai P.SD yaitu 0,66 untuk skala luas, sehingga diperoleh hasil pada Tabel 5.

Tabel 5. Hasil Uji Tingkat Kesukaran CPS Skala Luas

Kriteria	Kategori	Butir Soal
> 0,66	Sangat Sukar	S6, S10, S12
0 - 0,66	Sukar	S1, S4, S5, S13
- 0,66 - 0	Mudah	S3, S7, S9, S11, S14
< - 0,66	Sangat Mudah	S8, S2

Berdasarkan Tabel 5, tingkat kesukaran butir soal CPS pada uji coba skala luas menunjukkan distribusi yang bervariasi pada empat kategori, yaitu sangat sukar, sukar, mudah, dan sangat mudah. Sebagian butir soal berada pada kategori sukar dan mudah, sementara beberapa butir lainnya termasuk dalam kategori sangat sukar dan sangat mudah. Distribusi ini menunjukkan bahwa instrumen memiliki tingkat kesukaran yang seimbang dan tidak terpusat pada satu kategori tertentu, sehingga mampu mengukur kemampuan peserta didik pada berbagai tingkat secara lebih proporsional. Variasi tingkat kesukaran tersebut penting untuk memastikan bahwa instrumen tidak hanya mengukur peserta didik dengan kemampuan menengah, tetapi juga mampu mengidentifikasi peserta didik dengan kemampuan tinggi maupun rendah (Arikunto, 2011; Sumintono & Widhiarso, 2015). Dengan demikian, kualitas instrumen CPS dalam aspek tingkat kesukaran dapat dikatakan baik dan memenuhi kriteria soal yang efektif. Selanjutnya, uji reliabilitas instrumen CPS terhadap 66 responden menggunakan model Rasch dapat dilihat pada Gambar 7.

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	7.3	15.0	-.31	1.18				
SEM	.7	.0	.39	.07				
P.SD	5.9	.0	3.12	.53				
S.SD	6.0	.0	3.15	.54				
MAX.	15.0	15.0	4.46	1.86				
MIN.	.0	15.0	-4.39	.57				
REAL RMSE	1.31	TRUE SD	2.83	SEPARATION	2.15	Person	RELIABILITY	.82
MODEL RMSE	1.30	TRUE SD	2.84	SEPARATION	2.19	Person	RELIABILITY	.83
S.E. OF Person MEAN = .39								
Person RAW SCORE-TO-MEASURE CORRELATION = .99								
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .96 SEM = 1.17								
STANDARDIZED (50 ITEM) RELIABILITY = .94								

Gambar 7. Analisis Rasch Hasil Reliabilitas Uji Coba Luas Instrumen CPS

Berdasarkan Gambar 7, nilai Cronbach's alpha sebesar 0,96 menunjukkan bahwa instrumen CPS memiliki reliabilitas pada kategori sangat tinggi, sehingga hasil pengukuran dapat dianggap sangat konsisten dan dapat diandalkan (Sumintono & Widhiarso, 2015). Nilai separasi yang diperoleh sebesar 2,15 menunjukkan bahwa instrumen memiliki kemampuan yang sangat baik dalam membedakan tingkat kemampuan peserta didik. Hasil perhitungan nilai strata diperoleh sebesar 3,2 yang kemudian dibulatkan menjadi 3, sehingga butir soal dapat dikelompokkan ke dalam tiga kategori tingkat kesukaran, yaitu mudah, sedang, dan sulit. Hal ini menunjukkan bahwa instrumen memiliki kemampuan yang lebih baik dalam mengklasifikasikan tingkat kesukaran soal secara lebih rinci. Selanjutnya, hasil uji coba skala luas pada instrumen KPS dapat dilihat pada Gambar 8.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT MATCH OBS%	EXP%	Item	
7	30	66	.47	.42	1.07	.37	1.66	1.48	A	.76	.79	85.7	80.7	S7
13	32	66	.12	.43	1.58	2.15	1.12	.41	B	.74	.80	61.9	81.7	S13
8	38	66	-1.12	.49	1.52	1.65	1.07	.33	C	.79	.84	81.0	86.2	S8
1	36	66	-.67	.46	1.42	1.46	.93	.06	D	.79	.83	76.2	85.0	S1
9	36	66	-.67	.46	1.42	1.46	.93	.06	E	.79	.83	76.2	85.0	S9
6	20	66	2.06	.40	1.18	1.06	1.40	.72	F	.63	.67	71.4	76.3	S6
12	20	66	2.06	.40	1.18	1.06	1.40	.72	G	.63	.67	71.4	76.3	S12
5	28	66	.81	.41	.93	-.27	.99	.12	H	.77	.77	90.5	79.1	S5
10	30	66	-.47	.42	.96	-.11	.87	-.26	g	.79	.79	85.7	80.7	S10
2	42	66	-2.21	.56	.92	-.16	.65	.06	f	.84	.84	90.5	89.5	S2
11	36	66	-.67	.46	.84	-.52	.53	-.86	e	.86	.83	85.7	85.0	S11
14	32	66	.12	.43	.71	-1.25	.67	-.74	d	.84	.80	90.5	81.7	S14
3	34	66	-.26	.45	.45	-2.69	.27	-1.91	c	.85	.82	95.2	83.6	S3
4	34	66	-.26	.45	.45	-2.69	.27	-1.91	b	.85	.82	95.2	83.6	S4
15	34	66	-.26	.45	.45	-2.69	.27	-1.91	a	.85	.82	95.2	83.6	S15
MEAN	32.1	66.0	.00	.44	1.00	-.08	.87	-.28				83.5	82.5	
P.SD	5.8	.0	1.08	.04	.37	1.58	.42	1.07				9.8	3.5	

Gambar 8. Analisis Rasch Hasil Uji Coba Luas Instrumen KPS

Berdasarkan Gambar 8, seluruh 15 butir soal pada instrumen KPS dinyatakan valid, meskipun terdapat satu butir soal, yaitu nomor 7, yang hanya memenuhi dua dari tiga parameter yang ditetapkan, namun masih dapat dikategorikan valid sesuai kriteria yang digunakan. Hal ini menunjukkan bahwa secara umum butir soal telah memiliki kesesuaian yang baik dengan model Rasch, sehingga mampu merepresentasikan konstruk yang diukur secara memadai. Validitas yang tinggi pada seluruh butir soal mengindikasikan bahwa instrumen telah tersusun secara sistematis dan sesuai dengan indikator keterampilan proses sains yang dikembangkan. Selain itu, analisis tingkat kesukaran instrumen dilakukan dengan menggunakan nilai standar deviasi (SD) sebesar 1,08 sebagai dasar pengelompokan. Nilai tersebut kemudian digunakan untuk mengklasifikasikan butir soal ke dalam beberapa kategori tingkat kesukaran yang disajikan pada Tabel 6. Variasi tingkat kesukaran yang dihasilkan menunjukkan bahwa instrumen mampu mengakomodasi perbedaan kemampuan peserta didik secara lebih luas dan proporsional.

Tabel 6. Hasil Uji Tingkat Kesukaran KPS Skala Luas

Kriteria	Kategori	Butir Soal
> 1,08	Sangat Sukar	S6, S12
0 – 1,08	Sukar	S5, S7, S10, S13, S14
- 1,08 – 0	Mudah	S1, S3, S4, S9, S11, S15
< – 1,08	Sangat Mudah	S8, S2

Berdasarkan Tabel 6, tingkat kesukaran butir soal keterampilan proses sains (KPS) pada uji coba skala luas menunjukkan distribusi yang cukup beragam pada empat kategori, yaitu sangat sukar, sukar, mudah, dan sangat mudah. Sebagian besar butir soal berada pada kategori sukar dan mudah, yang menunjukkan bahwa instrumen dirancang dengan tingkat kesulitan yang seimbang untuk mengakomodasi kemampuan peserta didik pada level menengah. Sementara itu, keberadaan beberapa butir soal pada kategori sangat sukar dan sangat mudah memberikan variasi yang penting dalam mengidentifikasi peserta didik dengan kemampuan tinggi maupun rendah. Distribusi ini menunjukkan bahwa instrumen tidak terfokus pada satu tingkat kesukaran saja, melainkan mampu mengukur kemampuan peserta didik secara lebih komprehensif dan proporsional (Arikunto, 2011; Sumintono & Widhiarso, 2015). Dengan demikian, kualitas instrumen KPS dalam aspek tingkat kesukaran dapat dikatakan baik karena mampu membedakan tingkat kemampuan peserta didik secara lebih optimal. Selanjutnya, uji reliabilitas instrumen KPS menggunakan model Rasch dapat dilihat pada Gambar 9.

SUMMARY OF 66 MEASURED (EXTREME AND NON-EXTREME) Person								
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	6.5	14.0	-.40	1.14				
SEM	.7	.0	.34	.07				
P.SD	5.4	.0	2.74	.54				
S.SD	5.4	.0	2.76	.54				
MAX.	14.0	14.0	3.99	1.85				
MIN.	.0	14.0	-4.03	.56				
REAL RMSE	1.27	TRUE SD	2.43	SEPARATION	1.92	Person	RELIABILITY	.79
MODEL RMSE	1.26	TRUE SD	2.44	SEPARATION	1.94	Person	RELIABILITY	.79
S.E. OF Person MEAN = .34								
Person RAW SCORE-TO-MEASURE CORRELATION = .99								
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .95 SEM = 1.21								

Gambar 9. Analisis Rasch Hasil Reliabilitas Uji Coba Luas Instrumen KPS

Berdasarkan Gambar 9, nilai Cronbach's alpha sebesar 0,95 menunjukkan bahwa instrumen KPS memiliki reliabilitas pada kategori sangat tinggi, sehingga hasil pengukuran dapat dianggap sangat konsisten dan dapat diandalkan (Sumintono & Widhiarso, 2015). Nilai separasi yang diperoleh sebesar 1,92 menunjukkan bahwa instrumen memiliki kemampuan yang baik dalam membedakan tingkat kemampuan peserta didik. Hasil perhitungan nilai strata dengan menggunakan persamaan (1), diperoleh sebesar 2,9 yang kemudian dibulatkan menjadi 3, sehingga butir soal dapat dikelompokkan ke dalam tiga kategori tingkat kesukaran, yaitu mudah, sedang, dan sulit. Hal ini menunjukkan bahwa instrumen memiliki kemampuan yang lebih optimal dalam mengklasifikasikan tingkat kesukaran soal secara lebih rinci.

Hasil uji coba skala luas menunjukkan bahwa instrumen keterampilan proses sains (KPS) dan *creative problem solving* (CPS) yang dikembangkan telah memenuhi kriteria kelayakan secara komprehensif, baik dari aspek validitas, reliabilitas, daya beda, maupun tingkat kesukaran. Sebagian besar butir soal pada kedua instrumen dinyatakan valid berdasarkan parameter Rasch, dengan hanya beberapa butir yang tidak memenuhi seluruh kriteria namun tetap dapat diterima karena memenuhi sebagian parameter. Daya beda seluruh butir soal berada pada kategori tinggi, yang menunjukkan kemampuan instrumen dalam membedakan tingkat kemampuan peserta didik secara optimal. Selain itu, distribusi tingkat kesukaran yang bervariasi dari sangat mudah hingga sangat sukar menunjukkan bahwa instrumen mampu mengakomodasi berbagai tingkat kemampuan peserta didik secara proporsional. Nilai reliabilitas yang sangat tinggi pada kedua instrumen mengindikasikan bahwa hasil pengukuran bersifat konsisten dan dapat diandalkan. Dengan demikian,

instrumen yang dikembangkan tidak hanya layak digunakan dalam penilaian pembelajaran, tetapi juga memiliki kualitas yang baik dalam mengukur kemampuan KPS dan CPS secara lebih akurat dan mendalam.

Tahap Evaluation

Tahap *evaluation* dalam penelitian ini dilakukan secara berkelanjutan dengan mengintegrasikan hasil validasi teoritis oleh ahli dan hasil uji empiris menggunakan analisis Rasch. Perbaikan instrumen difokuskan pada tahap validasi ahli dengan mempertimbangkan masukan terkait aspek materi, konstruksi, dan bahasa. Hasil analisis empiris pada uji coba terbatas menunjukkan terdapat satu butir soal yang tidak valid, yaitu soal nomor 10 pada instrumen *creative problem solving* (CPS), karena tidak memenuhi parameter kesesuaian model. Namun, pada uji coba skala luas tidak ditemukan masukan substantif yang mengarah pada revisi konten, sehingga instrumen dipertahankan untuk menjaga konsistensi hasil pengukuran. Secara keseluruhan, diperoleh 14 butir soal CPS dan 15 butir soal keterampilan proses sains (KPS) yang dinyatakan valid. Nilai reliabilitas yang tinggi, yaitu 0,96 untuk CPS dan 0,95 untuk KPS, menunjukkan bahwa instrumen memiliki konsistensi yang sangat baik. Selain itu, variasi tingkat kesukaran dan daya beda yang terdistribusi secara proporsional menunjukkan bahwa instrumen mampu mengukur kemampuan peserta didik secara komprehensif, baik pada tingkat kemampuan rendah, sedang, maupun tinggi.

Pengembangan instrumen ini memberikan dampak positif terhadap praktik penilaian dalam pembelajaran fisika, khususnya dalam mengukur keterampilan proses sains dan kemampuan *creative problem solving* secara lebih terintegrasi. Instrumen yang dihasilkan memungkinkan guru untuk tidak hanya menilai penguasaan konsep, tetapi juga kemampuan berpikir ilmiah dan kreatif peserta didik dalam menyelesaikan permasalahan. Dengan adanya analisis berbasis Rasch, kualitas butir soal dapat dievaluasi secara objektif, sehingga hasil pengukuran menjadi lebih akurat dan dapat digunakan sebagai dasar dalam merancang strategi pembelajaran yang lebih tepat. Selain itu, instrumen ini juga dapat membantu guru dalam mengidentifikasi profil kemampuan peserta didik secara lebih rinci, sehingga mendukung implementasi pembelajaran yang berorientasi pada pengembangan keterampilan abad ke-21.

Meskipun demikian, penelitian ini masih memiliki beberapa kendala yang perlu diperhatikan. Instrumen yang dikembangkan masih terbatas pada aspek kognitif dan belum sepenuhnya mengakomodasi penilaian aspek afektif maupun psikomotorik yang juga penting dalam pembelajaran sains. Selain itu, bentuk instrumen yang masih berupa tes tertulis (pilihan ganda) menyebabkan keterbatasan dalam mengukur keterampilan proses sains secara langsung, yang pada dasarnya lebih optimal dinilai melalui observasi atau praktik. Penggunaan instrumen dalam bentuk cetak juga menjadi kendala dalam implementasi secara luas, terutama pada pembelajaran berbasis digital. Oleh karena itu, penelitian selanjutnya diharapkan dapat mengembangkan instrumen yang lebih komprehensif, terintegrasi dengan teknologi, serta mampu mengukur berbagai aspek kemampuan peserta didik secara lebih menyeluruh.

KESIMPULAN

Instrumen penilaian keterampilan proses sains (KPS) dan *creative problem solving* (CPS) yang dikembangkan dalam penelitian ini memberikan kontribusi penting dalam pengembangan asesmen berbasis *higher order thinking* melalui pendekatan Rasch yang mampu menghasilkan pengukuran yang objektif, valid, dan reliabel. Kebaruan penelitian ini terletak pada integrasi indikator KPS dan CPS dalam satu instrumen yang terstruktur serta penerapannya pada materi teori kinetik gas di jenjang SMA, sehingga mampu mengukur kemampuan berpikir ilmiah dan kreatif secara bersamaan. Instrumen ini tidak hanya berfungsi sebagai alat evaluasi, tetapi juga sebagai sarana diagnostik untuk membantu guru mengidentifikasi profil kemampuan peserta didik secara lebih komprehensif. Oleh karena itu, instrumen ini berpotensi menjadi alternatif dalam pengembangan sistem penilaian yang lebih bermakna dan kontekstual, serta mendorong praktik pembelajaran yang tidak hanya berorientasi pada hasil, tetapi juga pada proses berpikir. Ke depan, diperlukan pengembangan lebih lanjut dengan mengintegrasikan teknologi dan memperluas aspek penilaian agar instrumen semakin adaptif dan relevan dengan kebutuhan pembelajaran abad ke-21.

KONTRIBUSI PENULIS

Mira Fitri Yanti: Conceptualization, Methodology, Formal Analysis, Resources, Project Administration dan Writing - Original Draft; **Undang Rosidin:** Conceptualization, Methodology, dan Validation; serta **I Wayan Distrik:** Data Curation, Project Administration, dan Writing - Original Draft. Semua penulis telah membaca dan menyetujui versi akhir dari naskah ini.

PERNYATAAN KETERSEDIAAN DATA

Data yang mendukung temuan dalam penelitian ini tersedia dari para penulis atas permintaan yang wajar, dengan mempertimbangkan persetujuan etik dan peraturan institusi yang berlaku.

PERNYATAAN BEBAS KONFLIK KEPENTINGAN

Para penulis menyatakan bahwa tidak ada konflik kepentingan finansial maupun hubungan pribadi yang dapat mempengaruhi hasil yang dilaporkan dalam naskah ini.

PERNYATAAN ETIKA PENELITIAN DAN PUBLIKASI

Para penulis menyatakan bahwa penelitian dan penulisan naskah ini telah mematuhi standar etika penelitian dan publikasi, sesuai dengan prinsip ilmiah, serta bebas dari plagiasi.

PERNYATAAN PEMANFAATAN TEKNOLOGI ASISTIF

Para penulis menyatakan bahwa Kecerdasan Buatan Generatif (*Generative Artificial Intelligence*) dan teknologi asistif lainnya tidak digunakan secara berlebihan dalam proses penelitian dan penulisan naskah ini. Secara khusus, ChatGPT digunakan untuk *Brainstroming* ide, Grammarly untuk koreksi tata dan gaya bahasa, serta Scite untuk menganalisis konteks sitasi. Para penulis telah meninjau dan menyunting semua konten yang dihasilkan AI guna memastikan ketepatan, kelengkapan, serta kepatuhan terhadap standar etika dan ilmiah, dan bertanggung jawab penuh atas naskah versi akhir.

DAFTAR PUSTAKA

- Anjiana, R., Surahman, E., & Rizal, R. (2025). Urgensi scientific reasoning skills dan creative thinking skills dalam pendidikan: Analisis awal hasil peserta didik di sekolah. *DIFFRACTION: Journal for Physics Education and Applied Physics*, *7*(2), 83-93. Retrieved from: <https://jurnal.unsil.ac.id/index.php/Diffraction/article/view/16986>.
- Arikunto, S. (2011). *Prosedur penelitian suatu pendekatan praktik*. Jakarta: Rineka Cipta.
- Branch, R.M. (2009). *Instructional design-The ADDIE approach*. New York: Springer.
- Dinova, A. & Mustika, D. (2025). The relationship between creativity and the ability to solve HOTS science problems. *Inovasi Kurikulum*, *22*(4), 2199-2210. Retrieved from: <https://ejournal-hipkin.or.id/index.php/jik/article/view/181>.
- Dwikoranto, D., Lintangesukmanjaya, R.T., & Setiani, R. (2025). Analisis model rasch: Instrumen penilaian pemecahan masalah kreatif dalam pembelajaran fisika. *Proceeding Seminar Nasional IPA*, 507-514. Retrieved from: <https://proceedings.unnes.ac.id/snipa/article/view/4478>.
- Hanan, M.J., Irawan, M.P.T., & Naryatmojo, D.L. (2025). Evaluasi kesesuaian penilaian sumatif akhir jenjang kelas XII dengan distribusi ranah kognitif dalam kurikulum merdeka. *Jurnal Bastra (Bahasa dan Sastra)*, *10*(3), 912-922. Retrieved from: <https://bastra.uho.ac.id/index.php/journal/article/view/1504>.
- Hasan, Y., Syahrial, A., Busairi, A., & Doyan, A. (2024). Pengaruh strategi creative problem solving dengan pendekatan stem terhadap keterampilan proses sains dan kemampuan berpikir kreatif peserta didik. *Jurnal Ilmu Pendidikan dan Sosial (JIPSO)*, *1*(1), 22-28. Retrieved from: <https://ejournal.lembagaeinsteincollege.com/JIPSO/article/view/51>.
- 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), 1-17. DOI: <https://doi.org/10.1177/2158244020975409>.
- Khoiriyah, Z. & Shaleh, S. (2025). Solusi alternatif atas problematika dalam mengimplementasikan penilaian Higher Order Thingking Skills (HOTS). *Al-Madrasah: Jurnal Ilmiah Pendidikan Madrasah Ibtidaiyah*, *9*(2), 656-667. DOI: <http://dx.doi.org/10.35931/am.v9i2.4855>.
- Ling, M., Pang, V., & Salwana, B. (2014). Rasch analysis: Psychometric properties of Malaysian secondary school students' leadership scale. *Man In India*, *96*(1-2), 1-8. Retrieved from: https://serialsjournals.com/abstract/72216_1.pdf.
- Marjusa, P.P., Fatoni, A., & Hasanah, U. (2025). Efektivitas model creative problem solving terhadap kreativitas belajar pendidikan agama islam siswa SMP. *MODELING: Jurnal Program Studi PGMI*, *12*(4), 229-246. DOI: <https://doi.org/10.69896/modeling.v12i4.3035>.
- Matore, M.E.E.M., Maat, S.M., Affandi, H.M., Mohamad, S., & Khairani, A.Z. (2018). Assessment of psychometric properties for raven advanced progressive matrices in measuring Intellectual Quotient

- (IQ) using rasch model. *Asian Journal of Scientific Research*, **11**(3), 393-400. DOI: <https://doi.org/10.3923/ajsr.2018.393.400>.
- Nurhayati, N., Saputri, R.D., & Assegaf, A. (2019). Pengembangan instrumen keterampilan proses sains pada materi optika SMP. *Jurnal Edukasi*, **17**(2), 90–98. DOI: <https://doi.org/10.31571/edukasi.v17i2.1250>.
- Park, I.-S. & Kang, S.-H. (2012). The development of assessment tools to measure scientific creative problem solving ability for middle school students. *Journal of the Korean Association for Research in Science Education*, **32**(2), 210–223. DOI: <https://doi.org/10.14697/jkase.2012.32.2.21>.
- Purniasari, L., Masykuri, M., & Ariani, S.R.D. (2021). Analisis butir soal ujian sekolah mata pelajaran kimia SMA N 1 Kutowinangun tahun pelajaran 2019/2022 menggunakan model iteman dan rasch. *Jurnal Pendidikan Kimia*, **10**(2), 205-214. Retrieved from: <https://jurnal.uns.ac.id/JPKim/article/view/48244>.
- Putri, C.N. (2025). Pengaruh model creative problem solving pada pembelajaran matematika: Systematic literature review. *Jurnal PEKA (Pendidikan Matematika)*, **9**(1), 140-153. DOI: <https://doi.org/10.37150/qjyspx36>.
- Ramadhan, M.F., Siroj, R.A., & Afgani, M.W. (2024). Validitas and reliabilitas. *Journal on Education*, **6**(2), 10967-10975. DOI: <https://doi.org/10.31004/joe.v6i2.4885>.
- Riatma, A.T.A. & Susanti, A. (2025). Pengaruh model pembelajaran *creative problem solving* terhadap kemampuan pemecahan masalah matematis. *Hiradika: Jurnal Pendidikan dan Pembelajaran*, **2**(1), 8-13. Retrieved from: <https://jurnal.penerbithira.com/index.php/hiradika/article/view/53>.
- Setyaningrum, V. & Ikhtiyari, A. (2025). Identifikasi keterampilan proses sains di SMKN 1 Hanau: Diagnosa pada peserta didik baru. *Jurnal Tunas Pendidikan*, **8**(1), 174-183. DOI: <https://doi.org/10.52060/pgsd.v8i1.3427>.
- Sukma, N.K., Ismail, I., Faisal, F., Daud, F., & Taiyeb, M. (2026). Pengembangan modul belajar berbasis keterampilan proses sains (KPS) dalam meningkatkan hasil belajar pada materi perubahan lingkungan. *JiIP-Jurnal Ilmiah Ilmu Pendidikan*, **9**(2), 1830-1837. DOI: <https://doi.org/10.54371/jiip.v9i2.10581>.
- Sumintono & Widhiarso. (2015). *Aplikasi pemodelan RASCH pada assessment pendidikan*. Cimahi: Trim komunikata.
- Sund, R.B. & Trowbridge, L.W. (1973). *Teaching science by inquiry in the secondary school* (2nd ed.). Michigan: Merrill Publishing.
- Suryani, A., Siahaan, P., & Samsudin, A. (2015). Pengembangan Instrumen Tes untuk Mengukur Keterampilan Proses Sains Siswa SMP pada Materi Gerak. *Prosiding Simposium Nasional Inovasi dan Pembelajaran Sains 2015 (SNIPS 2015)*, 217-220. Retrieved from: https://ifory.id/proceedings/2015/z4pZjcJkq/snips_2015_ajeng_suryani_9471cb10c37438549b6ec72f59a49c2c.pdf.
- Susanti, D. (2026). Pendekatan pembelajaran matematika abad ke-21: Analisis literatur tentang berpikir kritis dan pemecahan masalah. *Jurnal Pendidikan Matematika*, **1**(1), 8-12. Retrieved from: <https://riset.nexusjurnal.com/index.php/JPM/article/view/45>.
- Treffinger, D.J. & Isaksen, S.G. (2005). Creative problem solving: History, development, and implications for gifted education and talent development. *Gifted Child Quarterly*, **49**(4), 342–353. DOI: <https://doi.org/10.1177/001698620504900407>.
- Wei, S., Liu, X., Wang, Z., & Wang, X. (2012). Using Rasch measurement to develop a compute modeling-based instrument to assess students' conceptual understanding of matter. *Journal of Chemical Education*, **89**(3), 335–345. DOI: <https://doi.org/10.1021/ed100852t>.
- Xiao, Y., Han, J., Koenig, K., Xiong, J., & Bao, L. (2018). Multilevel Rasch modeling of two tier multiple choice test: A case study using Lawson's classroom test of scientific reasoning. *Physical Review Physics Education Research*, **14**(2), 20104. DOI: <https://doi.org/10.1103/PhysRevPhysEducRes.14.020104>.
- Yurt, E. (2025). The creative problem-solving skills test: Development and initial validation. *International Journal of Education in Mathematics, Science and Technology*, **13**(3), 761–790. DOI: <https://doi.org/10.46328/ijemst.4711>.
- Zamri, A., & Nordin. (2015). Modeling a multiple choice mathematics test with the Rasch Model. *Indian Journal of Science and Technology*, **8**(12), 1–6. DOI: <https://doi.org/10.17485/ijst/2015/v8i12/70650>.