

Pengembangan Asesmen *High Order Thinking Skill* (HOTS) Terintegrasi Metakognitif Pada Materi Energi

Salsa Ayu Fajriyah¹, Woro Setyarsih²

^{1,2} Pendidikan Fisika, Universitas Negeri Surabaya, Surabaya, Indonesia

* Email: salsa.22084@mhs.unesa.ac.id

Abstrak

Kemampuan berpikir tingkat tinggi (*Higher Order Thinking Skills*/HOTS) merupakan salah satu kompetensi yang perlu dikembangkan dalam pembelajaran fisika. Namun, instrumen asesmen yang digunakan di sekolah umumnya masih berfokus pada hasil akhir dan belum mampu mengungkap proses berpikir siswa selama menyelesaikan masalah. Penelitian ini bertujuan mengembangkan instrumen asesmen HOTS terintegrasi *prompt* metakognitif pada materi energi dan hukum kekekalan energi, serta menganalisis karakteristik psikometri instrumen dan kemampuan HOTS siswa. Penelitian menggunakan metode *Research and Development* dengan model ADDIE yang meliputi tahap analisis, perancangan, pengembangan, implementasi, dan evaluasi. Tahap *Analyze* dilakukan melalui analisis kebutuhan, kurikulum, materi, dan kemampuan awal siswa yang menunjukkan perlunya instrumen yang dapat mengukur HOTS sekaligus memberikan informasi mengenai proses berpikir siswa. Berdasarkan hasil tersebut, tahap *Design* menghasilkan empat butir soal uraian pada level C4 dan C5 dengan komposisi tiga butir C4 dan satu butir C5, serta mengintegrasikan *prompt* metakognitif pada aspek *awareness*, *regulation*, dan *evaluation*. Tahap *Develop* meliputi validasi teoritis oleh ahli dan uji coba terbatas pada 32 mahasiswa Pendidikan Fisika untuk analisis menggunakan *Classical Test Theory*. Tahap *implement* melibatkan 82 siswa kelas X SMA yang dianalisis menggunakan *Rasch Model* berbantuan aplikasi Winsteps. Hasil penelitian menunjukkan bahwa instrumen yang dikembangkan memenuhi kriteria valid berdasarkan validasi ahli dan validitas empiris, serta memiliki reliabilitas, tingkat kesukaran, dan daya beda yang baik. Analisis Rasch menunjukkan seluruh butir soal memenuhi kriteria *fit*, nilai *item reliability* sebesar 0,90 dengan kategori baik, serta *person separation* sebesar 1,80 yang mampu membedakan kemampuan siswa ke dalam dua strata kemampuan siswa. Hasil implementasi juga menunjukkan bahwa kemampuan HOTS siswa didominasi pada kategori sedang. Integrasi *prompt* metakognitif tidak hanya menghasilkan instrumen yang memenuhi karakteristik psikometri yang baik, tetapi juga memberikan informasi mengenai proses berpikir siswa pada aspek *awareness*, *regulation*, dan *evaluation*. Dengan demikian, instrumen yang dikembangkan berpotensi dimanfaatkan sebagai alat asesmen sekaligus sarana diagnostik untuk mendukung perancangan pembelajaran fisika yang lebih efektif.

Kata kunci: Asesmen HOTS, Metakognitif, *Rasch Model*, Karakteristik Psikometri, Energi dan Hukum Kekekalan Energi, Kemampuan Berpikir Tingkat Tinggi

Development of a Metacognitively Integrated High-Order Thinking Skills (HOTS) Assessment for Energy-Related Content

Abstract

Higher-order thinking skills (HOTS) are essential competencies that should be fostered in physics learning. However, assessment instruments commonly used in schools tend to focus on students' final answers and have not adequately captured the thinking processes involved in problem solving. This study aimed to develop a HOTS assessment instrument integrated with metacognitive prompts on the topic of energy and the law of conservation of energy, as well as to examine its psychometric properties and students'

HOTS performance. This research employed a Research and Development approach using the ADDIE model, which consists of the analysis, design, development, implementation, and evaluation stages. The Analyze stage was conducted through an analysis of needs, curriculum, learning materials, and students' initial abilities, which indicated the need for an instrument capable of measuring HOTS while providing information about students' thinking processes. Based on the results, the Design stage produced four essay items at the C4 and C5 cognitive levels, consisting of three C4 items and one C5 item, with metacognitive prompts integrated into the aspects of awareness, regulation, and evaluation. The Develop stage included theoretical validation by three expert validators, while a limited trial involving 32 undergraduate physics education students was carried out to analyze the instrument using Classical Test Theory. The implementation stage involved 82 tenth-grade senior high school students, and the collected data were analyzed using the Rasch Model with Winsteps software. The results indicated that the developed instrument met the criteria for theoretical and empirical validity and demonstrated good reliability, item difficulty, and discrimination indices. Rasch analysis revealed that all items fit the measurement model, with an item reliability of 0.90, indicating high consistency, and a person separation index of 1.80, demonstrating the instrument's ability to classify students into two ability levels. The implementation results also showed that most students' HOTS were at the moderate level. Furthermore, integrating metacognitive prompts not only produced an assessment instrument with sound psychometric properties but also provided valuable information about students' metacognitive processes, including awareness, regulation, and evaluation. Therefore, the developed instrument has the potential to serve as both an assessment tool and a diagnostic instrument to support more effective physics instruction.

Keywords: *HOTS Assessment, Metacognitive, Rasch Model, Psychometric Characteristics, Energy and the Law of Conservation of Energy, Higher-Order Thinking Skill.*

Histori Naskah

Diserahkan: 27 Juli 2026

Direvisi: 8 Agustus 2026

Diterima: 12 Agustus 2026

How to cite:

Fajriyah S. A., Setyarsih W., (2026). Pengembangan Asesmen High Order Thinking Skill (HOTS) Terintegrasi Metakognitif Pada Materi Energi. *Inovasi Pendidikan Fisika*, 15(1), 37-49. DOI: <https://doi.org/10.26740/ipf.vxx.nxx.xx-xx>.

PENDAHULUAN

Perkembangan ilmu pendidikan di era abad ke-21 saat ini, tidak lagi cukup berfokus pada keterampilan penguasaan materi saja, namun harus membekali siswa dengan keterampilan yang dibutuhkan di era globalisasi yang mencakup kemampuan berpikir kritis, berkomunikasi secara efektif, bekerja sama dalam tim, serta berpikir kreatif dan inovatif yang dikenal dengan sebagai kompetensi *High Order Thinking Skill* (HOTS) (Dewangga & Sunarti, 2022). Sehingga pada peraturan menteri pendidikan dan kebudayaan nomor 20 tahun 2016 terkait Kurikulum 2016, peraturan baru disahkan sebagai bentuk paradigma baru dalam dunia pendidikan Indonesia yang mengarahkan pembelajaran sebagai upaya peningkatan keterampilan berpikir tingkat tinggi pada siswa (Purnasari et al., 2021). Dalam penelitiannya Parandreni et al., (2023) menyatakan bahwa pembelajarn dirancang dengan mengintegrasikan 3 hal penting yaitu, kurikulum, model/metode pembelajaran, dan asesmen. Oleh karena itu perangkat asesmen yang diterapkan harus sesuai dengan tujuan pembelajaran dalam meningkatkan kemampuan HOTS siswa.

Namun kenyataannya dalam pelaksanaan pembelajaran fisika yang membutuhkan kemampuan ilmiah yang logis dan terstruktur, justru seringkali instrumen asesmen yang digunakan hanya menguji kemampuan ingatan rumus dan pelaksanaan matematisnya tanpa menguji kemampuan HOTS siswa secara optimal (Givarin & Ellianawati, 2025). Beberapa penelitian menunjukkan bahwa sebagian besar asesmen yang

digunakan masih berpaku pada bentuk soal *Low Order Thinking Skill* (LOTS). Akibatnya, siswa kurang terlatih untuk memecahkan permasalahan soal non-rutin yang membutuhkan kemampuan bernalar, membenarkan argumen, atau menerapkan pengetahuan dalam konteks baru seperti soal-soal HOTS (Mor *et al.*, 2025). Akibatnya, siswa sering hanya mengejar jawaban akhir tanpa memahami proses pemecahan masalah yang sebenarnya.

Hasil survei internasional PISA (*Program for International Student Assessment*) tahun 2022 menunjukkan bahwa Indonesia berada di peringkat 69 dari 81 negara dengan skor literasi sains sebesar 383 poin, jauh di bawah rata-rata OECD (OECD, 2023). Rendahnya nilai capaian ini menunjukkan bahwa siswa masih mengalami kendala pada pengaplikasian konsep sains dalam memecahkan masalah kontekstual (Yusmar & Fadilah, 2023). Berdasarkan karakteristik soal PISA yang menuntut kemampuan siswa dalam analisis, penerapan konsep, serta evaluasi, ini menunjukkan bahwa keberhasilan siswa dalam menjawab soal dibutuhkan kemampuan dalam mengelola proses berpikir sebagai dukungan strategi metakognitif (T. Saputra *et al.*, 2025).

Penerapan asesmen yang berfokus pada kemampuan HOTS masih menghadapi berbagai kendala. Beberapa penelitian melaporkan bahwa sebagian instrumen yang diterapkan dalam pembelajaran masih didominasi oleh model soal rutin yang berorientasi pada LOTS (Dewangga & Sunarti, 2022). Sehingga seringkali bentuk soal fisika masih bersifat rutin dalam bentuk pilihan ganda yang hanya berfokus pada kemampuan mengingat hafalan rumus serta penyelesaiannya matematisnya (Masara *et al.*, 2024). Akibatnya siswa kurang terlatih dalam memecahkan permasalahan soal non-rutin yang membutuhkan kemampuan bernalar, penerapan konsep serta evaluasi argumen yang menunjukkan kemampuan HOTS dan kemampuan proses berpikir siswa (Mor *et al.*, 2025). Hal ini ditegaskan dengan hasil observasi di SMA Kartika IV-3 Surabaya yang menunjukkan bahwa hanya sebagian kecil siswa yang memiliki kemampuan berpikir tingkat tinggi sedangkan mayoritas siswa berada pada kategori kemampuan sedang hingga rendah, khususnya pada materi energi dan hukum kekekalan energi.

Kemampuan HOTS dalam menyelesaikan soal tidak hanya dipengaruhi oleh penguasaan konsep, tetapi juga dipengaruhi oleh kemampuan metakognitif dalam mengelola proses berpikir selama pemecahan masalah. Azlika *et al.*, (2025) menegaskan bahwa keterampilan metakognitif memiliki keterkaitan yang kuat dengan karakteristik soal HOTS karena siswa dituntut tidak hanya memahami konsep, tetapi juga berpikir secara reflektif, analitis, dan sadar terhadap langkah berpikir yang digunakan. Siswa yang mampu mengendalikan proses berpikirnya cenderung lebih berhasil menyelesaikan permasalahan kompleks dibandingkan siswa yang hanya mengandalkan hafalan konsep. Menurut Anderson & Krathwohl (2001), semakin tinggi level kognitif maka semakin dibutuhkan keterlibatan kemampuan proses metakognitif individu. Beberapa penelitian menjelaskan terkait metakognitif salah satunya pada penelitian Brown (1987) yang membagi metakognitif menjadi 2 komponen utama yaitu pengetahuan metakognitif dan regulasi metakognitif. Adapun Wilson & Clarke (2004) yang dalam penelitiannya menyebutkan bahwa regulasi metakognitif terdiri atas tiga komponen utama, yaitu *awareness*, *regulation*, dan *evaluation*, yang berfungsi membantu peserta didik menentukan strategi penyelesaian, memantau pelaksanaannya, serta mengevaluasi efektivitas hasil yang diperoleh.

Sejumlah penelitian terdahulu menunjukkan bahwa asesmen HOTS mampu mendorong siswa menghubungkan konsep dengan masalah kontekstual, namun penelitian yang mengembangkan instrumen asesmen secara khusus menintegrasikan metakognitif pada asesmen HOTS materi energi dan hukum kekekalan energi masih sangat terbatas (Ningsi & Shaleh, 2024). Penelitian oleh Nurmaliati *et al.* (2024) berfokus pada pengembangan instrumen HOTS fisika, sedangkan beberapa penelitian lain lebih menitikberatkan pada penerapan strategi metakognitif dalam pembelajaran. Sehingga dengan mengintegrasikan *prompt* tiap komponen metakognitif pada tiap butir soal sebagai *scaffolding*, diharapkan instrumen selain mampu mengukur hasil belajar namun juga dapat mengetahui cara berpikir siswa dalam strategi penyelesaian masalah sesuai komponen metakognitif (Noviansah, 2020). Selain itu, sebagian besar penelitian masih mengevaluasi kualitas instrumen hanya menggunakan analisis klasik sehingga informasi mengenai karakteristik psikometris butir soal, hubungan kemampuan siswa dengan tingkat kesukaran soal, serta kualitas pengukuran instrumen belum dianalisis secara mendalam menggunakan *Rasch Model* (Rahmatillah & Zainuddin, 2025).

Dengan demikian, penelitian ini diarahkan untuk menjawab bagaimana karakteristik psikometris instrumen asesmen HOTS terintegrasi metakognitif pada materi energi dan bagaimana profil HOTS metakognitif siswa. Analisis data dilakukan dengan dua metode sekaligus yaitu *Classical Test Theory* (CTT) dan *Rasch Model* untuk mendapatkan hasil analisis yang maksimal. Sehingga hasil penelitian diharapkan

menghasilkan instrumen asesmen yang berkualitas, mampu mendukung implementasi pembelajaran berbasis HOTS, serta memberikan alternatif asesmen yang tidak hanya berorientasi pada hasil akhir, tetapi juga pada proses berpikir peserta didik.

METODE

Penelitian ini menggunakan metode pendekatan *Research and Development* (RnD) dengan model ADDIE (*Analyze, Design, Develop, Implement, dan Evaluate*) (Branch, 2009). Model ADDIE di pilih karena tahapan penelitian pengembangannya sistematis dalam merancang, menguji, dan mengevaluasi produk instrumen yang dihasilkan.

Tahap *analyze* dilakukan dengan observasi terhadap objek awal permasalahan penelitian pengembangan melalui analisis kebutuhan, analisis kurikulum dan analisis materi. Pada tahap ini dilakukan observasi secara langsung melalui tes awal pada siswa dan wawancara guru untuk kemudian dievaluasi. Dari data diketahui bahwa pengembangan instrumen asesmen HOTS terintegrasi metakognitif diperlukan khususnya pada materi energi dan hukum kekekalan energi.

Tahap *design* dimulai dengan merancang instrumen asesmen, yang terdiri dari kisi-kisi instrumen, contoh bentuk soal, serta rubrik penilaiannya. Soal dirancang pada level kognitif menganalisis (C4) dan mengevaluasi (C5) dengan mengintegrasikan *prompt* metakognitif pada komponen *awareness, regulation, dan evaluation* sebagai *scaffolding* bagi siswa untuk menyadari proses berpikirnya. Butir soal yang dikembangkan terbatas pada jumlah 4 butir soal saja dengan setiap butir soalnya dilengkapi 3 tahap *prompt* metakognitif. Setiap tahapnya, terdapat satu hingga dua atau lebih soal *prompt* tambahan yang digunakan sebagai *scaffolding* untuk menuntun siswa dalam proses pemecahan masalah. Siswa diarahkan untuk terlebih dahulu mengidentifikasi informasi penting pada narasi soal sehingga siswa mampu memahami pengetahuan awal yang relevan dengan permasalahan untuk dilakukan tahap selanjutnya, hal ini didasarkan pada tahap komponen *awareness*. Selanjutnya, siswa pada tahap *regulation* diarahkan untuk merencanakan strategi penyelesaian masalah yang sesuai narasi stimulus soal. Tahap terakhir dilakukan dengan siswa diminta mengulas kembali jawaban yang diperoleh pada tahap-tahap sebelumnya. Adapun dari empat butir soal yang dikembangkan, tiga butir dirancang untuk mengukur kemampuan C4 dan satu butir untuk mengukur kemampuan C5. Komposisi tersebut disesuaikan dengan karakteristik dari materi energi yang menuntut siswa dalam menganalisis informasi, hubungan antar besaran serta konsep yang relevan sebagai kemampuan menganalisis (C4). Kemampuan mengevaluasi (C5) juga relevan dengan materi energi yang menuntut siswa untuk menilai ketepatan hasil atau strategi penyelesaian berdasarkan konsep dan informasi yang tersedia. Sedangkan kemampuan mengkerasi (C6) tidak dibuat karena karakteristik tuntutan kognitif pada materi dan permasalahan yang dikembangkan tidak mengarahkan siswa untuk menghasilkan atau merancang suatu produk, gagasan, maupun alternatif penyelesaian baru. Sehingga perancangan instrumen hanya difokuskan pada ranah kognitif C4 dan C5 yang sesuai dengan karakteristik materi yang di kembangkan. Adapun kisi-kisi yang dirancang dapat diketahui dari tabel berikut.

Tabel 1. Rancangan Kisi-Kisi Instrumen

Tujuan Pembelajaran	Indikator Butir Soal	Sub Bab Materi	Level Kognitif	Bentuk Soal	No. Soal
Memahami bentuk-bentuk energi yaitu energi kinetik, energi potensial dan energi mekanik serta memahami hukum kekekalan energi mekanik dalam permasalahan sehari-hari.	Peserta didik dapat menganalisis hubungan antara perubahan energi potensial dengan usaha yang dilakukan mesin untuk menyimpulkan alasan ketidakmampuan kendaraan dalam meningkatkan kecepatan saat membawa beban maksimal.	Energi Mekanik	C4 (Menganalisis)	Essay	1
	Peserta didik dapat mengevaluasi efektivitas sistem keamanan menggunakan konsep usaha dan energi, serta menyusun rekomendasi perbaikan secara ilmiah pada sistem <i>crumple zone</i> .	Perubahan Energi	C5 (Mengevaluasi)	Essay	2

Tujuan Pembelajaran	Indikator Butir Soal	Sub Bab Materi	Level Kognitif	Bentuk Soal	No. Soal
	Peserta didik mampu menganalisis proses transformasi energi kinetik dan potensial serta membuktikan hukum kekekalan energi mekanik pada titik tengah lintasan melalui data tabel.	Energi potensial dan energi kinetik	C4 (Menganalisis)	Essay	3
	Peserta didik mampu menganalisis perubahan energi mekanik pada titik A, B, dan C serta menghitung kecepatan kereta di titik B dan C menggunakan hukum kekekalan energi untuk menentukan apakah kereta mampu mencapai titik C	Energi Mekanik	C4 (Menganalisis)	Essay	4

Tahap *develop* melanjutkan hasil rancangan instrumen asesmen dengan menguji validasi teoritis instrumen yang dilakukan melalui validasi ahli dan uji validasi secara empiris, lalu uji reliabilitas, tingkat kesukaran, dan daya beda butir soal menggunakan pendekatan *Classical Test Theory* (CTT) setelah pelaksanaan ujicoba terbatas yang dilakukan pada 32 mahasiswa prodi Pendidikan Fisika Universitas Negeri Surabaya. Untuk kemudian hasil analisisnya digunakan untuk mengevaluasi tiap butir soal dan menentukan kualitas butir soal sebelum dilanjut pada tahap implementasi.

Tahap *implement* diterapkan pada siswa kelas X SMA Kartika IV-3 Surabaya sebanyak 82 siswa yang sebelumnya telah mendapatkan materi ajar energi dan hukum kekekalan energi. Kemudian hasil jawaban siswa yang didata dan dianalisis digunakan untuk mengetahui karakteristik psikometri instrumen serta analisis kemampuan HOTS siswa melalui pendekatan *Rasch Model* berbantuan aplikasi Winstep. Analisis validitas butir soal menggunakan data *output item measure* dengan mengikuti syarat batas *outfit Mean Square* (MNSQ) yang diterima sekitar $0,5 < \text{MNSQ} < 1,5$, batas nilai *outfit Z-Standard* (ZTSD) yang diterima sebesar $-2,0 < \text{ZTSD} < +2,0$ dan batas nilai *Point Measure Correlation* (*Pt Measure Corr*) yang diterima sekitar $0,4 < \text{Pt Measure Corr} < 0,85$ (Rahmatillah & Zainuddin, 2025). Butir soal dikatakan valid saat butir soal telah memenuhi sedikitnya dua syarat. Selain itu, karakteristik butir soal pada reliabilitas berdasarkan *rasch model* memiliki kriteria berikut.

Tabel 2. Tabel Kriteria Reliabilitas Pada *Rasch Model*

Nilai Reliabilitas (<i>Person/Item</i>)	Interpretasi
$> 0,94$	Istimewa
$0,91 - 0,94$	Bagus Sekali
$0,81 - 0,90$	Bagus
$0,67 - 0,80$	Cukup
$< 0,67$	Lemah

Selanjutnya untuk menginterpretasikan kriteria taraf kesukaran butir soal digunakan data *output* yang sama pada data validitas. Pengkategorian tingkat kesukaran berdasarkan nilai *measure logit* dan nilai simpangan baku logit item yang diperoleh (Cahyani et al., 2024). Untuk penginterpretasian digunakan empat kriteria taraf kesukaran dari data berikut.

Tabel 3. Kriteria Taraf Kesukaran *Rasch Model*

Taraf kesukaran (<i>Measure logit</i>)	Kriteria
$\text{Measure logit} < -\text{SD logit}$	Sangat Mudah
$-\text{SD logit} \leq \text{Measure logit} \leq 0$	Mudah
$0 < \text{Measure logit} \leq \text{SD logit}$	Sulit
$\text{Measure logit} > \text{SD logit}$	Sangat Sulit

Sedangkan daya beda dalam pendekatan *Rasch Model*, data yang dianalisis adalah tabel *output item-person measure* untuk mendapatkan nilai separation menggunakan rumus berikut.

$$H = \frac{(4 \times \text{Separation}) + 1}{3}$$

Tahap *evaluate* dilakukan di setiap akhir pelaksanaan pertahap pengembangan yang telah dilakukan. Dengan pelaksanaan tahap evaluasi di tiap tahapan pengembangan, maka diharapkan dapat menghasilkan produk instrumen asesmen yang memenuhi kriteria dari karakteristik psikometri instrumen, sehingga layak untuk digunakan sebagai alat ukur kemampuan HOTS siswa.

HASIL DAN PEMBAHASAN

Hasil Pengembangan

Hasil produk dari penelitian pengembangan ini adalah instrumen asesmen HOTS terintegrasi metakognitif dengan tipe soal uraian. Soal terdiri dari 4 butir soal yang tiap butir soalnya memiliki *prompt* dari ketiga komponen metakognitif yang diperuntukkan sebagai *scaffolding* bagi siswa untuk mengetahui cara berpikir yang tepat dalam menyelesaikan masalah dalam soal. Materi yang digunakan adalah materi energi dan hukum kekekalan energi untuk siswa kelas X SMA/Sederajat.

Instrumen yang dikembangkan dalam penelitian ini melalui lima tahapan pengembangan ADDIE (Branch, 2009). Tahap analisis digunakan untuk menganalisis tingkat keurgensian penelitian sekaligus observasi lapangan secara langsung di SMA Kartika IV-3 Surabaya. Selanjutnya instrumen dirancang sesuai dengan tujuan penelitian dengan merancang, kisi-kisi, contoh bentuk soal, dan rubrik penilaian. Bentuk soal dengan integrasi metakognitif dapat dilihat pada tabel berikut.

Tabel 4. Contoh Bentuk Instrumen Asesmen Yang Dikembangkan

Keterangan	Bentuk Soal
Stimulus dan soal utama asesmen HOTS	1. Pada suatu perjalanan karya wisata di kota Pacet, sekolah SMA Kartika IV-3 Surabaya menyewa sebuah bus pariwisata dengan berisikan 35 penumpang dalam satu armada bus ukuran sedang. Perjalanan arah Pacet yang menanjak dan berkelok membuat mesin bus harus bekerja lebih keras. Saat bus melaju ke pusat wisata yang berada di ketinggian 6 meter dari titik awal, sopir menambahkan tekanan pada pedal gas agar bus mendapatkan energi tambahan untuk bus tetap melaju di jalan yang menanjak. Walau begitu kecepatan bus sulit bertambah. Jika massa bus kosong sebesar 5000 kg sedangkan massa tiap penumpang beserta bawaannya sebesar 60 kg. Tentukan energi tambahan yang dibutuhkan mesin bus untuk mencapai pusat wisata. Selanjutnya analisislah bagaimana perubahan energi yang terjadi memengaruhi beban kerja mesin, serta jelaskan mengapa kecepatan bus tidak bertambah meskipun energi mesin meningkat! (Asumsikan percepatan gravitasi adalah 10 m/s^2)
Integrasi <i>prompt</i> metakognitif indikator <i>awareness</i>	Awareness (Sebelum mengerjakan, lengkapi peta pikiran berikut) a. Informasi apa saja yang menunjukkan massa dan perubahan posisi sistem pada situasi tersebut? b. Agar kendaraan mampu bergerak menuju tempat yang lebih tinggi maka dibutuhkan energi tambahan apa?
Integrasi <i>prompt</i> metakognitif indikator <i>regulation</i>	Regulation (Tuliskan strategi penyelesaian Anda) a. Rumus apa yang kamu gunakan untuk menghitung energi tambahan yang dibutuhkan bus, dan bagaimana langkah perhitungannya? b. Bagaimana cara kamu menghubungkan hasil perhitungan dengan konsep perubahan energi mekanik dalam sistem kendaraan?
Integrasi <i>prompt</i> metakognitif indikator <i>evaluation</i>	Evaluation (Setelah mendapatkan jawaban angka, periksa kembali) a. Jelaskan kembali bagaimana hasil perhitungannya, terkait perubahan energi yang memengaruhi kerja mesin dan mengapa kecepatan bus tidak bertambah meskipun mesin menghasilkan energi lebih besar!

Adapun pada tahap *development*, instrumen dilakukan uji validitas ahli dengan bantuan 3 validator ahli untuk menguji validitas instrumen berdasarkan aspek ranah isi, konstruk dan bahasa (Saputra et al., 2024). Instrumen yang valid secara teoritis kemudian dilakukan ujicoba untuk kemudian dianalisis kualitas instrumen berdasarkan validitas, reliabilitas, tingkat kesukaran dan daya beda butir soal dengan pendekatan *Classical Test Theory* (CTT). Hasil baca analisis kualitas produk awal instrumen kemudian digunakan untuk merevisi atau menentukan butir soal yang dapat digunakan atau tidak sebagai produk utama.

Selanjutnya instrumen yang telah ditetapkan sebagai produk utama, dilanjutkan pada tahap implementasi dengan menguji luaskan pada siswa SMA Kartika IV-3 Surabaya sebanyak 82 siswa. Dari data

ini kemudian didapatkan informasi terkait karakteristik psikometri instrumen dan kemampuan HOTS siswa melalui model *Rasch Model* berbantuan aplikasi Winstep.

Hasil Uji Validitas Teoritis

Validasi ahli dilakukan oleh 3 validator ahli yaitu, 2 dosen Prodi Pendidikan Fisika dan 1 guru Fisika SMA. Validasi dilakukan dengan memberikan lembar instrumen kepada validator ahli untuk dinilai kevalidan instrumen berdasarkan 3 ranah penilaian yaitu isi, konstruk, dan bahasa. Yang kemudian data yang didapatkan, dianalisis menggunakan formula *Percentage of Agreement* (PoA). Data hasil PoA yang didapatkan, dapat dilihat pada tabel 5.

Tabel 5. Hasil Uji Validasi Ahli

Ranah Penilaian	Persentase PoA	Kriteria
Isi	100,00%	Valid
Konstruk	95,70%	Valid
Bahasa	85,71%	Valid

Berdasarkan hasil data Tabel 5 diketahui bahwa instrumen asesmen HOTS terintegrasi metakognitif memiliki nilai persentase PoA tiap ranah penilaian diatas persentase minimal >75% kriteria kelayakan (Borich, 1994). Sehingga instrumen dapat dinyatakan valid dan layak digunakan. Walau begitu saran perbaikan tetap ada terkait perumusan bahasa dan kejelasan struktur pertanyaan untuk di revisi sebelum digunakan untuk pelaksanaan uji coba. Adapun hasil revisi dapat ditunjukkan pada Tabel 6.

Tabel 6. Hasil Revisi Berdasarkan Saran Validator Ahli

Saran dan Revisi	Nomor Butir Soal
Merevisi tipografi dan kesalahan tanda baca	1,2,3,4
Menyesuaikan kalimat yang lebih efektif dari "Perjalanan arah Pacet yang melalui daerah bukit berkelok membutuhkan tenaga ekstra bus" menjadi "Perjalanan arah Pacet yang menanjak dan berkelok membuat mesin bus harus bekerja lebih keras"	1
Mengganti kalimat "Tragedi duka kecelakaan kereta api di Stasiun Bekasi Timur menjadi duka mendalam bagi banyak orang" menjadi " Dalam sebuah kecelakaan antar kereta api di Stasiun Bekasi Timur yang melibatkan antara KRL yang sedang berhenti darurat di rel yang sama dengan KA Argo Bromo Anggrek"	2
Memberikan kalimat cerita yang lebih praktis dengan mengganti kalimat "seorang petani buah kelapa sedang melakukan pemanenan dan seorang siswa sedang mengamati selama proses pemanenean" menjadi " Seorang siswa sedang mengamati proses pemanenan buah kelapa"	3
Memberikan cetak miring pada kata asing seperti "roller coaster" menjadi <i>roller coaster</i>	4

Hasil Uji Coba

Setelah tahap evaluasi dan revisi selesai, uji validitas teoritis dilakukan melalui tahap uji coba terbatas dan didapatkan hasil analisis yang menunjukkan beberapa informasi terkait kualitas instrumen yang termasuk kedalam hasil validitas empiris, tingkat kesukaran butir soal, daya beda, dan reliabilitas instrumen. Data validitas empiris ditunjukkan pada Tabel 7.

Tabel 7. Hasil Uji Coba Validitas Empiris

No Butir Soal	$r_{hitung} (r_{xy})$	r_{tabel}	Kriteria
1	0,726	0,349	Tingkat Valid Tinggi

No Butir Soal	$r_{hitung} (r_{xy})$	r_{tabel}	Kriteria
2	0,807		Tingkat Valid Tinggi
3	0,767		Tingkat Valid Tinggi
4	0,787		Tingkat Valid Tinggi

Hasil data validitas empiris menunjukkan bahwa instrumen memiliki kriteria valid tinggi dengan nilai $r_{hitung} (r_{xy})$ tiap butir soal yang lebih tinggi dari nilai r_{tabel} instrumennya. Semakin tinggi tingkat kriteria kevalidan instrumen, maka instrumen semakin menunjukkan keakuratan dalam mengukur data (Rahayu et al., 2024). Adapun data tingkat kesukaran butir soal didapatkan dan ditunjukkan pada Tabel 8.

Tabel 8. Hasil Uji Coba Tingkat Kesukaran Butir Soal

No Butir Soal	Nilai	Kriteria
1	0,656	Sedang
2	0,726	Mudah
3	0,648	Sedang
4	0,508	Sedang

Hasil uji tingkat kesukaran butir soal diketahui bahwa dari keempat butir soal terdapat tiga butir soal dengan tingkat kesukaran sedang dalam kisaran nilai 0,31-0,70 dan satu butir soal dengan tingkat kesukaran mudah dalam kisaran nilai 0,71-1,00. Sehingga instrumen tidak memiliki satupun butir soal dengan tingkat kesukaran sulit. Selanjutnya untuk hasil uji daya beda didapatkan data sebagai berikut.

Tabel 9. Hasil Uji Coba Daya Beda

No Butir Soal	Nilai	Kriteria
1	0,56	Baik
2	0,50	Baik
3	0,50	Baik
4	0,67	Baik

Hasil daya beda diketahui bahwa keempat butir soal instrumen memiliki tingkat daya beda yang baik. Data yang ditunjukkan pada Tabel 9 melihat bahwa keseluruhan nilai daya beda butir soal berada di kisaran $0,40 < DB \leq 0,70$ dengan kriteria baik.

Tabel 10. Hasil Uji Coba Reliabilitas

Cronbach Alpha (r_{xy})	N
0,76882	4

Nilai reliabilitas instrumen yang bernilai lebih dari nilai min *alpha cronbach* sehingga instrumen dapat dinyatakan reliabel. Sehingga instrumen menunjukkan konsistensi walau pengukuran dilakukan berulang kali (Widaningrum et al., 2023). Dengan hasil analisis dari data uji coba yang sudah didapatkan, instrumen di evaluasi untuk mendapatkan butir soal instrumen yang layak digunakan sebagai produk utama hasil pengembangan instrumen ini. Berdasarkan informasi data yang sudah didapatkan maka, diketahui bahwa dari keempat instrumen layak digunakan setelah dilakukan evaluasi dan revisi.

Karakteristik Psikometri Instrumen

Uji sebar luas dilakukan pada tahap *implement* yang kemudiannya datanya diolah menggunakan bantuan aplikasi winstep untuk mengetahui karakteristik psikometri instrumen dengan baik. Analisis karakteristik psikometri instrumen terdiri dari analisis validitas (*fit*), tingkat kesukaran, daya beda dan reliabilitas yang dianalisis dari *output item measure*.

Tabel 11. *Output Item Measure*

No	Tingkat Kesukaran	Outfit MNSQ	Outfit ZTSD	PTMC
4	0,50	0,81	-1,30	0,72
3	0,32	0,99	0,00	0,62
1	-0,26	1,08	0,60	0,65
2	-0,57	0,99	0,00	0,67

Validitas butir soal dianalisis berdasarkan Tabel 11 dengan menilai kecocokan butir soal dengan kriteria pengukuran berdasarkan batas nilai $0,5 < \text{MNSQ} < 1,5$ untuk *outfit* MNSQ, $-2,0 < \text{ZTSD} < +2,0$ untuk *outfit* ZSTD, dan nilai $0,4 < \text{Pt Measure Corr} < 0,85$. Sehingga diketahui bahwa seluruh butir soal dikatakan valid (*fit*). Adapun tingkat kesukaran butir soal dilihat dari urutan nomor paling kanan. Dengan nomor butir soal 4 sebagai *item* butir dengan kesukaran sangat sulit dan untuk nomor butir soal paling bawah yaitu nomor 2 menjadi *item* butir soal dengan tingkat kesukaran sangat mudah.

Tabel 12. *Output Item-Person Measure*

Indikator	Nilai
<i>Item Reliability</i>	0,90
<i>Person Separation</i>	1,80

Untuk mengetahui informasi daya beda dan reliabilitas instrumen, analisis dilakukan pada *output item-person measured*. Data *item reliability* untuk menunjukkan konsistensi tingkat kesukaran butir soal dan *person separation* menunjukkan kemampuan instrumen dalam memisahkan tingkat kemampuan yang dimiliki siswa (Bond & Fox, 2015). Berdasarkan Tabel 12, nilai *item reliability* menunjukkan bahwa reliabilitas instrumen berada pada kriteria bagus. Sedangkan nilai *person separation* dengan nilai 1,80 menunjukkan bahwa instrumen mampu membagi tingkat kemampuan siswas berdasarkan 2 strata pembagian tingkat kemampuan siswa.

Hasil keseluruhan uji karakteristik psikometri instrumen berdasarkan *rasch model* menunjukkan bahwa instrumen yang dikembangkan dalam penelitian memiliki kualitas instrumen asesmen yang baik. Seluruh butir soal memenuhi kriteria syarat kevalidan butir soal (*fit*). Sehingga setiap butir soal dikatakan mampu mengukur kemampuan secara konsisten. Adapun variasi tingkat kesukaran yang diperoleh pada butir soal terbagi menjadi 4 menurut kriteria pada Tabel 11 yaitu sangat mudah, mudah, sulit, dan sangat sulit. Sehingga keempat butir soal memiliki tingkat kesukaran yang berbeda. Adapun nilai reliabilitas yang bagus dan daya beda yang baik sehingga cukup mampu membedakan kemampuan siswa. Sehingga dengan hasil analisa tersebut, menurut Saputra et al., (2024) instrumen telah memenuhi persyaratan sebagai alat ukur yang berkualitas, sehingga layak digunakan untuk mengevaluasi kemampuan peserta didik serta memberikan informasi yang akurat mengenai tingkat kesukaran butir dan distribusi kemampuan responden.

Kemampuan HOTS Metakognitif Siswa

Penerapan analisis *Rasch Model* menggunakan aplikasi winstep memungkinkan evaluasi terhadap butir soal terkait karakteristik instrumen asesmen dan kemampuan responden (*person ability*) (Rahmatillah & Zainuddin, 2025). Dalam penelitian ini, untuk mengetahui kemampuan HOTS metakognitif siswa berdasarkan uji luas, maka data di analisis berdasarkan data *output person measure* yang hasil pengelompokannya dapat diketahui sebagai berikut.

Tabel 13. *Output Person Measure*

Kriteria Kemampuan	Nilai Logit	Jumlah <i>person</i>
Tinggi	2,29	7
	1,72	8
Sedang	1,27	5
	0,88	13
	0,54	6
	0,22	10
	-0,08	13
	-0,37	4
	-0,66	5
Rendah	-0,66	5

Kriteria Kemampuan	Nilai Logit	Jumlah person
	-0,97	5
	-1,30	2
	-1,69	4

Mengetahui dari Tabel 13, tingkat kemampuan siswa dibagi menjadi 3 yaitu tingkat kemampuan tinggi, sedang, dan rendah. Kemampuan siswa dibagi berdasarkan nilai logit tiap *person* sehingga diketahui bahwa tingkat kemampuan tinggi terdiri dari dua skala logit, tingkat kemampuan sedang terdiri dari enam logit, dan tingkat kemampuan rendah dengan jumlah empat skala logit. Pembagian ini mengikuti hasil perhitungan nilai mean dan standar deviasi sebagai ketentuan nilai batas skala logit yang dapat ditunjukkan pada tabel berikut.

Tabel 14. Kriteria Kemampuan Siswa

Logit Person	Kemampuan Siswa
$x > 1,45$	Tinggi
$1,45 \geq x \geq -0,65$	Sedang
$x < -0,65$	Rendah

Sehingga berdasarkan Tabel 14, diketahui bahwa siswa dengan kemampuan tingkat tinggi yang berjumlah 15 *person* memiliki nilai logit di atas 1,45 logit. Sedangkan pada 51 siswa dengan tingkat kemampuan sedang, memiliki nilai logit di antara -0,65 logit hingga 1,45 logit. Adapun bagi 16 siswa dengan tingkat kemampuan rendah, mereka memiliki nilai logit dibawah -0,65 logit.

Dengan jumlah sebaran nilai logit dan tingkat kemampuan siswa tersebut. Dapat diketahui bahwa kemampuan keterampilan berpikir tingkat tinggi siswa dalam menyelesaikan instrumen asesmen HOTS terintegrasi metakognitif tergolong cukup baik. Mayoritas siswa berada di kategori sedang hingga rendah. Data mean person pada skala 0,40 logit dan mean item pada skala 0,00 logit menunjukkan adanya selisih jarak logit yang mengindikasikan bahwa kemampuan siswa jauh lebih tinggi dalam menyelesaikan butir soal daripada tingkat kesukaran butir soal. Semakin tinggi nilai *person measure* maka semakin tinggi tingkat kemampuan siswa dalam mengerjakan soal (Sumintono & Widhiarso, 2015). Hal ini menunjukkan bahwa pengintegrasian *prompt scaffolding* metakognitif dalam butir soal telah sedikit banyak berhasil membantu siswa dalam mengetahui cara berpikir dalam menyelesaikan permasalahan soal.

Awareness (Sebelum mengerjakan, lengkapi peta pikiran berikut)

a. Informasi apa saja yang menunjukkan massa dan perubahan posisi sistem pada situasi tersebut?
 Massa bus kosong = 5.000 kg, M 35 penumpang beserta barang = 35 x 60 = 2.100 kg
 m total sistem = 5.000 + 2.100 = 7.100 kg

b. Agar kendaraan mampu bergerak menuju tempat yang lebih tinggi maka dibutuhkan energi tambahan apa?
 Energi potensial gravitasi, yaitu energi yg diperlukan untuk mengangkat bus beserta penumpangnya ke tempat yg lebih tinggi.

Regulation (Tuliskan strategi penyelesaian Anda)

a. Rumus apa yang kamu gunakan untuk menghitung energi tambahan yang dibutuhkan bus, dan bagaimana langkah perhitungannya?

Rumus E. Potensial
 $E_p = mgh$
 $m = 5000 + (35 \times 60)$
 $m = 5000 + 2100 = 7100 \text{ kg}$

2. Hitung E. Tambahan
 $E_p = 7100 \times 10 \times 6$
 $E_p = 426000 \text{ J}$

Bagaimana cara kamu menghubungkan hasil perhitungan dengan konsep perubahan energi mekanik dalam sistem kendaraan?
 Hasil perhitungan menunjukkan bahwa energi tambahan sebesar 426.000 J digunakan untuk menaikkan energi potensial gravitasi saat bus naik ke tempat yg lebih tinggi, energi dari mesin berubah menjadi energi potensial gravitasi bus. Bus juga akan bertambah beratnya.

Evaluation (Setelah mendapatkan jawaban angka, periksa kembali)

1. Jelaskan kembali bagaimana hasil perhitungannya, terkait perubahan energi yang memengaruhi kerja mesin dan mengapa kecepatan bus tidak bertambah meskipun mesin menghasilkan energi lebih besar!
 Bus membutuhkan energi tambahan sebesar 426.000 J untuk mencapai ketinggian 6 m. Energi dari mesin digunakan untuk melawan gaya gravitasi dan menghasilkan energi potensial bus, karena sebagian besar energi yg di pakai untuk melawan, hanya sedikit energi yg tersisa utk meningkatkan energi kinetik.

Gambar 1. Hasil Jawaban Siswa dengan Kemampuan Tinggi

Hasil jawaban siswa dengan kemampuan tinggi menunjukkan kemampuan metakognitif yang lebih matang, dengan mampu merencanakan, mengontrol, dan mengevaluasi proses berpikir secara mandiri. Kolom tahap *awareness* membuktikan kemampuan siswa menemukan informasi data yang di butuhkan serta mampu memahami permasalahan soal. Adapun pada tahap *regulation*, mampu menyelesaikan perhitungan yang dibutuhkan serta mampu menjelaskan keterhubungan hasil perhitungannya dengan konsep hukum mekanik. Sama halnya pada tahap *evaluation*, siswa mampu dengan baik menganalisis keterhubungan hasil perhitungan perubahan energi dengan beban kerja mesin bus. Hal ini sejalan dengan penelitian Silaban & Darhim (2023) juga menunjukkan bahwa siswa dengan kemampuan metakognitif yang baik cenderung memiliki kemampuan berpikir tingkat tinggi yang lebih baik karena mampu mengontrol proses berpikirnya secara sadar. Berbeda dengan jawaban siswa berkemampuan rendah dan sedang.

Penyelesaian:

1. A. bus keluar setiap 5 ton
Hasil penumpang dengan bendungan
varian 1445 = $40 \times 35 + 1000$

• B. energi potensial = 7100 kg

Penyelesaian

A. $EP = m \cdot g \cdot h$
 $= 7100 \cdot 10 \cdot 6$
 $= 7100 \cdot 60$
 $EP = 426000 \text{ J}$

B. jika energi : Potensial Mula
Mula energi : Kinetik Mula

Evaluasi

Apakah saat bus naik maka berat dari bus akan meningkat
karena terdapat daya gravitasi bumi

1,2 a.) 35 penumpang
diket: h: 6 meter
m: 500 kg
g: 10 m/s
ep: 7100, 10.6
I: 426, 000 //

a.) Siswa dengan Kemampuan Sedang

b.) Siswa dengan Kemampuan Rendah

Gambar 2. Hasil Jawaban Siswa dengan Kemampuan Sedang dan Rendah

Berdasarkan contoh Gambar 2 diketahui bahwa jawaban siswa dengan kemampuan sedang umumnya mampu menjawab seluruh soal atau setengah keseluruhan soal dari semua tahapan *prompt* metakognitif secara tepat. Sehingga beberapa dari siswa masih mengalami kegagalan memahami konsep terhadap permasalahan kontekstual yang disajikan dalam soal. Sedangkan bagi sebagian besar siswa dengan kemampuan rendah justru hanya mampu menyelesaikan secara tepat pada tahap *awareness* dan gagal menyelesaikan tahap *prompt* metakognitif selanjutnya. Azlika et al., (2025) dalam penelitiannya menyebutkan bahwa siswa dengan kemampuan metakognitif sedang dalam menyelesaikan soal HOTS umumnya hanya mampu menyelesaikan 2 indikator komponen metakognitif.

Sehingga dengan dapat diketahuinya kemampuan metakognitif siswa melalui pengintegrasian *prompt* metakognitif dalam asesmen HOTS, proses asesmen tidak hanya difokuskan pada capaian hasil belajar saja. Tetapi juga mampu menunjukkan bagaimana siswa memahami permasalahan (*awareness*), memilih strategi (*regulation*), serta mengevaluasi kembali hasil yang didapatkan (*evaluation*). Dengan begitu, instrumen asesmen dapat dipergunakan sebagai sarana diagnostik yang memberikan informasi yang dapat dimanfaatkan bagi guru untuk merancang pembelajaran yang lebih tepat sasaran dengan menyesuaikan kebutuhan dan karakteristik proses berpikir siswa.

KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, instrumen asesmen HOTS terintegrasi metakognitif pada materi energi dan hukum kekekalan energi menunjukkan karakteristik psikometris yang baik berdasarkan analisis *Classical Test Theory* (CTT) dan *Rasch Model*. Keempat butir soal dinyatakan valid dan reliabel, memiliki daya beda yang baik, serta memenuhi kriteria *fit* pada analisis Rasch dengan nilai *item reliability* sebesar 0,90 dan *person separation* sebesar 1,80, dengan tingkat kesukaran butir yang bervariasi dari sangat mudah hingga sangat sulit. Adapun kemampuan HOTS metakognitif siswa menunjukkan bahwa dari 82 siswa, sebanyak 15 siswa berada pada kategori tinggi, 51 siswa pada kategori sedang, dan 16 siswa pada kategori rendah, sehingga kemampuan siswa secara umum didominasi kategori sedang. Integrasi *prompt* metakognitif juga menunjukkan adanya perbedaan proses berpikir antarkategori kemampuan, yaitu siswa berkemampuan tinggi mampu menunjukkan *awareness*, *regulation*, dan *evaluation* secara lebih lengkap, sedangkan siswa berkemampuan sedang dan rendah masih mengalami kendala terutama pada tahap *regulation* dan *evaluation*. Dengan demikian, instrumen yang dikembangkan tidak hanya memiliki karakteristik psikometris yang baik sebagai alat ukur HOTS, tetapi juga mampu memberikan gambaran mengenai profil kemampuan HOTS-metakognitif siswa melalui proses *awareness*, *regulation*, dan *evaluation*.

KONTRIBUSI PENULIS

Salsa Ayu Fajriyah: Conceptualization, Methodology, Investigation, Data Curation, Formal Analysis, Writing – Original Draft; serta **Woro Setyarsih:** Conceptualization, Methodology, Validation, Writing – Review & Editing, Supervision. Semua penulis telah membaca dan menyetujui versi akhir dari naskah ini.

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 membantu penyusunan struktur naskah, penyempurnaan tata bahasa, penyuntingan kalimat agar lebih efektif dan sesuai dengan kaidah bahasa Indonesia ilmiah, serta memberikan masukan dalam penyusunan uraian pembahasan berdasarkan data yang telah diperoleh penulis. Para penulis telah meninjau, memverifikasi, dan menyunting seluruh konten yang dihasilkan oleh AI guna memastikan ketepatan, kesesuaian dengan data penelitian, kelengkapan, serta kepatuhan terhadap standar etika dan ilmiah, dan bertanggung jawab penuh atas naskah versi akhir.

DAFTAR PUSTAKA

- Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.
- Azlika, L., Aklmawati, Mahmuzah, R., Fajriana, & Zahara, Y. (2025). Analisis Kemampuan Metakognisi Siswa dalam Menyelesaikan Soal High Order Thinking Skill (HOTS). *Jurnal Pendidikan Matematika Malikussaleh*, 5(1), 1–8. <https://doi.org/https://doi.org/10.29103/jpmm.v5i1.17613>
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences (3rd Edition)*. Routledge.
- Borich, G. D. (1994). *Observation Skill for Effective Teaching*. New York: Mc.Graw Hill Companies.
- Branch, R. M. (2009). *Instructional Design: The ADDIE Approach*. New York: Springe.
- Brown, A. L. (1987). *Metacognition, Motivation, and Understanding*. Lawrence Erlbaum Associates Publishers.
- Cahyani, N. A. P., Suryana, D., & Sunarya, Y. (2024). Pengembangan Instrumen Citra Tubuh: Analisis Rasch Model. *Jurnal Psikologi*, 20(2), 132–145. <https://doi.org/http://dx.doi.org/10.24014/jp.v14i2.21041>
- Dewangga, N. S. A., & Sunarti, T. (2022). Analisis Soal Fisika Berbasis High Order Thinking Skill dalam Penilaian Akhir Tahun di SMA. *Pendipa Journal of Science Education*, 6(2), 573–579. <https://doi.org/https://doi.org/10.33369/pendipa.6.2.573-579> ABSTRACT
- Givarin, D., & Ellianawati. (2025). Analisis Pembuatan Soal HOTS untuk Mengembangkan Kemampuan Berpikir Tingkat Tinggi Siswa dalam Pemecahan Masalah Fisika. *Jurnal Sosial Humaniora Dan Pendidikan: Scripta Humanika*, 1(1), 1–7. <https://scriptaintelektual.com/scripta-humanika/index>
- Masara, A., Gunantiani, J. M., & Rohman, S. M. (2024). Analysis of Students' Difficulties in Understanding Black's Principle in Physics Education. *Jurnal Pendidikan Dan Ilmu Fisika (JPiF)*, 04(02), 205–214. <https://doi.org/https://doi.org/10.52434/jpif.v4i2.1583>
- Mor, B., Patel, R. N., & Prajapati, B. (2025). 21st-Century Competencies in Physics: Assessment Strategies for Critical Thinking, Problem-Solving, and Character Formation. *Schrödinger: Journal of Physics Education*, 6(2), 135–145. <https://doi.org/10.37251/sjpe.v6i2.2030>
- Ningsi, A., & Shaleh. (2024). Problematika Penerapan Asesmen Berbasis HOTS (Higher Order Thinking Skills) di Sekolah Dasar. *SAP (Susunan Artikel Pendidikan)*, 8(3), 447–455. <https://doi.org/https://doi.org/10.30998/sap.v8i3.20998>
- Noviansah, A. (2020). Objek dalam Assesment Penilaian (Afektif, Kognitif, dan Psikomotorik). *Al-Hikmah Jurnal Studi Islam*, 1(2), 114–124. <https://ejournal.kopertais4.or.id/sasambo/index.php/alhikmah/article/view/3832>
- Parandengi, R., Anwar, L., & Sa'dijah, C. (2023). Pengembangan Asesmen yang Mendukung Kemampuan Pemecahan Masalah Matematis Siswa Berbasis Project-Based Instruction. *Aksioma: Jurnal Program Studi Pendidikan Matematika*, 12(1), 459–472. <https://doi.org/https://doi.org/10.24127/ajpm.v12i1.6264>
- Purnasari, P. D., Silvester, & Lumbantobing, W. L. (2021). Pengembangan Instrumen Asesmen Higher Order Thingking Skills (HOTS) Ditinjau dari Gaya Belajar Siswa. *Sebatik*, 25(2), 571–580. <https://doi.org/10.46984/sebatik.v25i2.1607>
- Rahayu, Y., Ruhiat, Y., & Atikah, C. (2024). Analisis Validitas, Reliabilitas, dan Tingkat Kesukaran Instrumen Tes Pemahaman Konsep Trigonometri pada Siswa Kelas XI. *Jurnal Teknologi Pendidikan Dan Pembelajaran*, 11(2), 27–36. <https://doi.org/https://dx.doi.org/10.62870/jtppm.v11i2>

- Rahmatillah, N., & Zainuddin, A. (2025). Evaluation of HOTS Test on Renewable Energy Materials Through Rasch Model Analysis. *Jurnal Pijar MIPA*, 20(5), 933–939. <https://doi.org/https://doi.org/10.29303/jpm.v20i5.9458>
- Republik, I. K. P. dan K. (2016). *Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 20 Tahun 2016 tentang Standar Kompetensi Lulusan Pendidikan Dasar dan Menengah*. <https://peraturan.bpk.go.id/Details/224177/permendikbud-no-20-tahun-2016>
- Saputra, O. A., Nurfajriani, & Sudrajat, A. (2024). Analisis Instrumen Penilaian Hasil Belajar Kimia untuk Mengukur High Order Thinking Skill (HOTS) Semester Genap Kelas X Berdasarkan Permodelan Rasch. *Didaktika: Jurnal Kependidikan*, 13(2), 1967–1984. <https://doi.org/10.58230/27454312.730>
- Saputra, T., Purnomo, E. A., & Joko, I. (2025). Peningkatan Kemampuan Higher Order Thinking Skills (HOTS) Melalui Optimalisasi Metakognitif Siswa: A Systematic Literature Review. *Jurnal Ilmiah Matematika (JIMAT)*, 6(1), 396–408. <https://doi.org/https://doi.org/10.63976/jimat.v6i1.854> e-ISSN
- Silaban, B., & Darhim. (2023). Kemampuan Metakognisi Siswa Berdasarkan Gaya Kognitif dalam Menyelesaikan Soal PISA. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, 07(02), 1496–1507. <https://doi.org/https://doi.org/10.31004/cendekia.v7i2.2242>
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Model Rasch untuk Penelitian Ilmu Sosial*. Jakarta: Tim Komunikata Publishing House.
- Widaningrum, D., Mindyarto, B. N., & Aji, M. P. (2023). Development of an Assessment as Learning Based on an ISCACoRe Metacognitive Strategy to Improve Problem- Solving Skills. *Jurnal Penelitian Pendidikan IPA (JPPIPA)*, 9(5), 2548–2555. <https://doi.org/10.29303/jppipa.v9i5.2441>
- Wilson, J., & Clarke, D. (2004). Towards the Modelling of Mathematical Metacognition. *Mathematics Education Research Journal*, 16(2), 25–48.
- Yusmar, F., & Fadilah, R. E. (2023). Analisis Rendahnya Literasi Sains Peserta Didik Indonesia: Hasil PISA dan Faktor Penyebab. *Lensa (Lentera Sains): Jurnal Pendidikan IPA*, 13(1), 11–19. <https://doi.org/10.24929/lensa.v13i1.283>