



Standarisasi Tes Keterampilan Berpikir Kreatif Kinematika

Muh. Tawil^{1✉}, Haksan Darwangsa²

Universitas Negeri Makassar, Indonesia¹, Keguruan dan Ilmu Pendidikan Universitas Mulawarman,
Indonesia²

e-mail : muh.tawil@unm.ac.id¹, haksandarwangsa2019@gmail.com²

Abstrak

Latar belakang penelitian ini adalah tes keterampilan berpikir kreatif di abad 21 sangat urgen diteliti karena kompetensi penting yang harus dimiliki siswa agar mampu menghadapi perkembangan ilmu pengetahuan dan teknologi. Penelitian ini bertujuan untuk menstandarisasi tes keterampilan berpikir tinggi (KBT) kinematika peserta didik kelas IX SMA Negeri di Kabupaten Gowa meliputi aspek: (1) validitas, (2) reliabilitas, (3) tingkat kesukaran, dan (4) daya pembeda. Penelitian bersifat deskriptif kuantitatif dengan populasi 251 peserta didik dari 7 kelas. Sampel ditentukan menggunakan rumus Slovin sebanyak 154 peserta didik melalui teknik random sampling. Data dikumpulkan dengan tes pilihan ganda sebanyak 20 butir soal dan dianalisis menggunakan Microsoft Excel. Hasil penelitian menunjukkan: (1) instrumen valid dengan nilai koefisien konsistensi internal VR = 1 (tinggi); (2) instrumen reliabel R = 100%; (3) tingkat kesukaran dengan kategori mudah rata-rata 20,5%; kategori sedang rata-rata 52,87%, dan sukar rata-rata 87,5%; (4) daya pembeda sangat baik (ID = 0,50), cukup baik (ID = 0,41), kurang baik (ID = 0,41), dan tidak baik (ID = 0,07). Secara keseluruhan, tes KBT memenuhi syarat standarisasi. Secara keseluruhan, tes KBT memenuhi syarat standarisasi.

Kata Kunci: keterampilan; berpikir tingkat tinggi; standarisasi; tes

Abstract

The background of this research is that the creative thinking skill test in the 21st century is very urgent to be researched because of the important competencies that students must have to be able to face the development of science and technology. This study aims to standardize a kinematics-based Higher-Order Thinking Skills (HOTS) test for 9th-grade students at a public high school in Gowa Regency, covering the following aspects: (1) validity, (2) reliability, (3) difficulty level, and (4) discriminating power. This descriptive-quantitative study involved a population of 251 students across seven classes. A sample of 154 students was selected using Slovin's formula via random sampling. Data were collected using a 20-item multiple-choice test and analyzed using Microsoft Excel. The results indicate that: (1) the instrument is valid, with an internal consistency coefficient (VR) of 1 (high); (2) the instrument is reliable, with an R-value of 100%; (3) regarding difficulty levels, the average distribution was 20.5% for "easy," 52.87% for "medium," and 87.5% for "difficult"; and (4) discriminating power values ranged from very good (ID = 0.50) and quite good (ID = 0.41) to poor (ID = 0.41) and very poor (ID = 0.07). Overall, the HOTS test meets the requirements for standardization.

Keywords: skills; higher-order thinking; standardization; test.

Copyright (c) 2026 Muh. Tawil, Haksan Darwangsa

✉ Corresponding author :

Email : muh.tawil@unm.ac.id

DOI : <https://doi.org/10.31004/edukatif.v8i5.9412>

ISSN 2656-8063 (Media Cetak)

ISSN 2656-8071 (Media Online)

PENDAHULUAN

Keterampilan berpikir kreatif merupakan salah satu kompetensi utama abad ke-21 yang harus dimiliki peserta didik agar mampu menghadapi perkembangan ilmu pengetahuan, teknologi, dan tantangan kehidupan nyata. Dalam pembelajaran fisika, keterampilan berpikir kreatif berperan penting dalam menghasilkan berbagai alternatif penyelesaian masalah, mengembangkan ide-ide baru, serta menghubungkan konsep fisika dengan fenomena kehidupan sehari-hari (Noviani dkk., 2025). OECD bahkan menempatkan *creative thinking* sebagai salah satu kompetensi yang diukur dalam PISA karena kemampuan tersebut berkontribusi terhadap keberhasilan peserta didik dalam memecahkan masalah yang kompleks (OECD, 2024).

Salah satu materi fisika yang memerlukan keterampilan berpikir kreatif adalah kinematika. Materi ini tidak hanya menuntut kemampuan menghitung besaran gerak, tetapi juga kemampuan menginterpretasikan grafik, menganalisis fenomena gerak, serta menghasilkan berbagai alternatif penyelesaian terhadap permasalahan kontekstual. Namun, evaluasi pembelajaran fisika di sekolah masih didominasi oleh soal-soal yang mengukur kemampuan mengingat konsep dan penyelesaian matematis rutin sehingga belum mampu menangkap kemampuan berpikir kreatif peserta didik secara optimal.

Beberapa penelitian telah mengembangkan instrumen keterampilan berpikir kreatif dalam pembelajaran fisika. Santofani & Rosana, (2016) mengembangkan tes kreativitas pada materi teori kinetik gas berbasis pendekatan inkuiri yang memenuhi kriteria validitas, reliabilitas, dan tingkat kesukaran yang baik. Selanjutnya, Tanjung & Nasution, (2022) mengembangkan instrumen keterampilan berpikir kreatif pada materi listrik arus searah berdasarkan indikator Torrance yang menghasilkan instrumen valid dan reliabel. Penelitian lain oleh Utama dkk. (2021) juga mengembangkan instrumen keterampilan berpikir kreatif berbasis *Project Based Learning* untuk pembelajaran fisika yang berfokus pada validasi isi dan konstruk. Selain itu, Kurniawan dkk. (2025) menunjukkan bahwa penggunaan analisis Aiken's V dan Rasch Model mampu menghasilkan instrumen tes berpikir kreatif fisika yang memiliki validitas isi, reliabilitas, dan karakteristik butir yang baik.

Meskipun demikian, penelitian mengenai instrumen keterampilan berpikir kreatif pada materi kinematika masih relatif terbatas. Penelitian Tiadarma, Sahyar, dan Bunawan, (2022) telah berhasil mengembangkan instrumen keterampilan berpikir kreatif pada materi kinematika dan memperoleh karakteristik butir yang baik. Akan tetapi, penelitian tersebut lebih menitikberatkan pada tahap pengembangan instrumen dan belum menghasilkan instrumen baku yang dilengkapi proses standarisasi secara komprehensif sehingga dapat digunakan secara luas sebagai alat ukur kemampuan berpikir kreatif peserta didik.

Berdasarkan kondisi tersebut, diperlukan penelitian mengenai standarisasi tes keterampilan berpikir kreatif pada materi kinematika. Standarisasi menjadi tahapan penting karena tidak hanya memastikan validitas dan reliabilitas instrumen, tetapi juga menghasilkan karakteristik butir, tingkat kesukaran, daya pembeda, fungsi informasi, serta norma interpretasi skor yang dapat digunakan secara konsisten pada berbagai kelompok peserta didik. Instrumen yang terstandarisasi diharapkan mampu memberikan hasil pengukuran yang objektif dan akurat sehingga dapat dimanfaatkan sebagai dasar pengambilan keputusan dalam pembelajaran maupun penelitian pendidikan fisika.

Kebaruan (*novelty*) penelitian ini terletak pada proses standarisasi tes keterampilan berpikir kreatif yang secara khusus dikembangkan pada materi kinematika. Berbeda dengan penelitian terdahulu yang umumnya hanya mengembangkan dan memvalidasi instrumen, penelitian ini menekankan penyusunan instrumen baku melalui analisis psikometrik yang lebih komprehensif sehingga menghasilkan perangkat tes yang memenuhi persyaratan sebagai instrumen standar untuk mengukur keterampilan berpikir kreatif peserta didik pada materi kinematika.

METODE

Objek Penelitian

Penelitian ini menggunakan literature sebanyak 28 artikel yang memenuhi standar jurnal terindeks garuda dan scopus. Artikel ini digunakan pada uraian pendahuluan, review literature dan pembahasan. Tahapan atau prosedur yang digunakan dalam penelitian, yakni pertama menentukan objek penelitian dalam hal ini peserta didik kelas IX SMAN di Kabupaten Gowa Sulawesi Selatan yang telah mempelajari kinematika sebanyak 7 (tujuh) kelas dengan jumlah populasi $N = 251$ peserta didik. Penentuan jumlah sampel menggunakan rumus Slovin dengan tingkat kesalahan 5%, sehingga diperoleh sebanyak 154 peserta didik sebagai sampel penelitian. Teknik pengambilan sampel menggunakan simple random sampling, sehingga setiap anggota populasi memiliki peluang yang sama untuk terpilih sebagai sampel.

Jenis dan Sumber Data Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan jenis penelitian deskriptif. Penelitian bertujuan untuk mendeskripsikan hasil standarisasi tes keterampilan berpikir kreatif (KBK) pada materi kinematika melalui analisis karakteristik instrumen tes. Penelitian dilaksanakan pada peserta didik kelas IX tingkat SMAN yang telah mempelajari materi kinematika.

Berdasarkan dari pelaksanaan penelitian ini, maka diperoleh data berupa data deskriptif, yakni skor hasil validasi dari pakar pengetahuan alam dan pakar pendidikan IPA; dan hasil tes keterampilan berpikir kreatif. Semua data tersebut dianalisis secara kuantitatif dan kualitatif.

Partisipan Penelitian

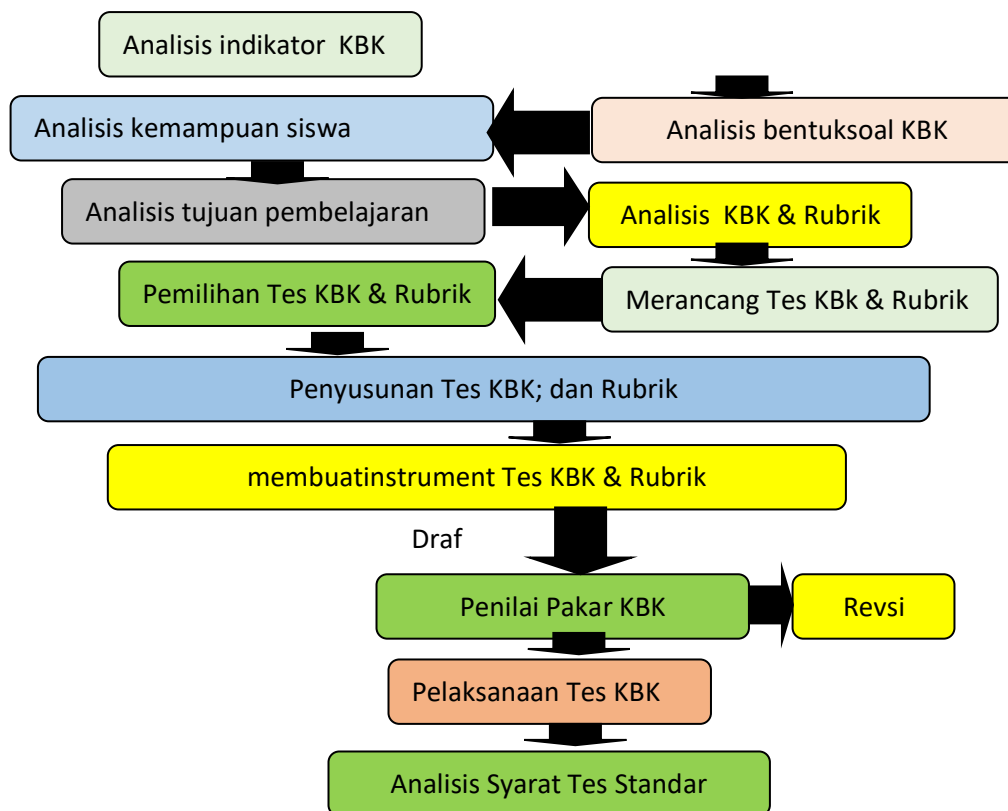
Partisipan dalam penelitian ini terdiri dari berbagai pihak yang memiliki peran penting dalam standarisasi tes KBK ini. Mereka meliputi pakar ilmu pengetahuan alam (IPA) dan pakar pendidikan IPA, kepala sekolah SMN se kabupaten Gowa, para guru fisika; ketua lembaga penelitian dan pengabdian masyarakat (LP2M) yang telah mendanai penelitian ini; tenaga observer, dan serta para peserta didik kelas IX SMAN se kabupaten Gowa sebagai subjek penelitian.

Proses dan Teknik Pengumpulan Data

Proses pengumpulan data dalam penelitian ini dilakukan melalui 4 (empat) teknik utama: (1) pengembangan instrument KBK. Instrumen penelitian berupa tes pilihan ganda sebanyak 20 butir soal yang disusun berdasarkan indikator KBK pada materi kinematika. Sebelum digunakan dalam pengambilan data, instrumen disusun berdasarkan kisi-kisi yang mengacu pada capaian pembelajaran dan indikator KBK. Validasi instrumen, oleh ahli (expert judgment) untuk menilai kesesuaian isi, konstruksi, dan bahasa. Selanjutnya, instrumen diujicobakan kepada responden untuk memperoleh data empiris yang digunakan dalam proses standarisasi; dan analisis data, meliputi valid, reliable, daya pembeda dan tingkat kesukaran item. Pengumpulan data dilakukan melalui pemberian tes kepada seluruh sampel penelitian. Jawaban peserta didik diberi skor 1 untuk jawaban benar dan 0 untuk jawaban salah. Data hasil tes kemudian direkap dalam lembar kerja Microsoft Excel untuk dianalisis. Tahap berikutnya; (2) revisi. Kegiatan ini bersifat siklik hingga dihasilkan produk tes KBK yang valid dan reliable; (3) analisis koefisien konsistensi internal dan reliabilitas tes KBK; daya beda, dan tingkat kesukaran soal.

Teknik Analisis dan Keabsahan Data

Proses analisis data kevalidan menggunakan analisis koefisien konsistensi Gregory (Borich, 2015; Gregory, 2004; Pradana, 2025; Ramdani et al., 2021; Tawil & Liliarsari, 2025). Kriteria AKPS perangkat, dan instrumennya reliable jika nilai reliabilitasnya (R) $\geq 0,75$ dan Analisis peningkatan persentasi rata-rata skor menggunakan n-gain atau normal gain (Ramdani et al., 2021; Tawil & Liliarsari, 2025) menghasilkan interpretasi ilmiah yang bermanfaat bagi pengembangan teori dan praktik pendidikan Islam di era digital (Azman et al., 2025). Untuk mencapai tujuan ini, digunakan teori pengembangan Four-D dengan. Rancangan penelitian dapat dilihat pada Gambar 1.



Gambar 1. Diagram Alir Penelitian

HASIL DAN PEMBAHASAN

Validitas Tes KBK

Hasil validasi oleh 2 validator ahli terhadap instrumen tes KBK menunjukkan bahwa komponen validasi konstruksi soal dan validasi isi soal masing-masing memperoleh nilai koefisien konsistensi internal $VR = 1$ (sangat tinggi). Hasil ini menunjukkan bahwa tes KBT yang dikembangkan telah memenuhi kesesuaian antara butir soal dengan indikatornya serta selaras dengan tujuan pembelajaran yang ingin diukur.

Hasil Analisis Validasi Tes Keterampilan Berpikir Kreatif (KBT) Materi Kinematika

Validasi instrumen tes KBK materi kinematika dilakukan melalui penilaian oleh para ahli menggunakan analisis koefisien konsistensi internal Gregory. Validasi ini bertujuan untuk mengetahui tingkat kesesuaian instrumen terhadap aspek konstruk dan isi sebelum diujicobakan kepada peserta didik. Hasil analisis validasi disajikan pada Tabel 1.

Tabel 1. Hasil Analisis Validasi Tes KBK

No.	Komponen	Koefisien Relevansi	Kategori
1	Validasi konstruk soal KBK	1,00	Sangat tinggi
2	Validasi isi soal KBK	1,00	Sangat tinggi

Berdasarkan hasil analisis pada Tabel 1, diperoleh nilai koefisien relevansi sebesar 1,00 pada aspek validasi konstruk dan validasi isi. Nilai tersebut termasuk dalam kategori sangat tinggi, yang menunjukkan bahwa seluruh butir soal memiliki tingkat kesepakatan yang sempurna antarvalidator. Dengan demikian, setiap butir soal dinilai telah sesuai dengan indikator keterampilan berpikir tingkat tinggi, tujuan pembelajaran, karakteristik materi kinematika, serta kaidah penyusunan instrumen evaluasi.

Pada aspek validasi konstruk, koefisien Gregory sebesar 1,00 mengindikasikan bahwa seluruh butir soal telah memenuhi aspek konstruksi yang baik, meliputi kejelasan rumusan soal, kesesuaian stimulus dengan indikator, penggunaan bahasa yang komunikatif, serta ketepatan penyusunan alternatif jawaban. Hal ini menunjukkan bahwa instrumen telah dirancang sesuai dengan prinsip penyusunan tes yang mampu mengukur keterampilan berpikir tingkat tinggi peserta didik.

Sementara itu, pada aspek validasi isi, koefisien sebesar 1,00 menunjukkan bahwa materi yang diukur dalam setiap butir soal telah sesuai dengan kompetensi yang ingin dicapai pada materi kinematika. Seluruh indikator keterampilan berpikir tingkat tinggi, seperti kemampuan menganalisis, mengevaluasi, dan mencipta, telah terwakili dalam instrumen sehingga tes memiliki kesesuaian isi (content validity) yang sangat baik.

Secara keseluruhan, hasil validasi menggunakan koefisien Gregory menunjukkan bahwa instrumen tes keterampilan berpikir tingkat tinggi materi kinematika memiliki validitas isi dan validitas konstruk yang sangat tinggi. Oleh karena itu, instrumen dinyatakan layak untuk dilanjutkan pada tahap uji coba empiris guna menganalisis karakteristik butir soal, meliputi validitas empiris, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh sebelum digunakan sebagai instrumen standar dalam mengukur keterampilan berpikir tingkat tinggi peserta didik kelas XI SMA Negeri.

Reliabilitas Tes KBK

Hasil Analisis Reliabilitas Tes Keterampilan Berpikir Kreatif (KBK)

Reliabilitas instrumen tes KBK pada materi kinematika dianalisis menggunakan Percentage of Agreement (PoA). Analisis ini bertujuan untuk mengetahui tingkat konsistensi penilaian antarvalidator terhadap setiap butir soal. Dalam analisis PoA, reliabilitas ditentukan berdasarkan nilai rata-rata agreement. $\bar{d}(A)$, yaitu tingkat kesepakatan antarvalidator, dan rata-rata disagreement. $\bar{d}(D)$, yaitu tingkat ketidaksepakatan antarvalidator. Semakin tinggi nilai *agreement* dan semakin rendah nilai *disagreement*, maka semakin tinggi pula reliabilitas instrumen.

Berdasarkan hasil analisis pada Tabel 2, seluruh butir soal memperoleh nilai rata-rata disagreement ($\bar{d}(D)$) sebesar 0, sedangkan nilai rata-rata agreement ($\bar{d}(A)$) berada pada rentang 3,0 hingga 4,0. Hasil tersebut menghasilkan nilai **reliabilitas (R)** sebesar 100% untuk seluruh 20 butir soal yang dianalisis.

Nilai reliabilitas sebesar 100% menunjukkan bahwa seluruh validator memiliki tingkat kesepakatan yang sempurna dalam memberikan penilaian terhadap setiap butir soal. Tidak adanya nilai *disagreement* mengindikasikan bahwa tidak terdapat perbedaan pendapat di antara validator mengenai kualitas setiap butir soal, baik ditinjau dari aspek kesesuaian indikator, kejelasan konstruksi, penggunaan bahasa, maupun kesesuaian materi yang diukur.

Variasi nilai $\bar{d}(A)$ yang berkisar antara 3,0; 3,5; hingga 4,0 menunjukkan adanya perbedaan jumlah skor kesepakatan yang diberikan oleh validator pada masing-masing butir soal. Namun demikian, karena seluruh validator memberikan penilaian yang konsisten tanpa adanya ketidaksepakatan ($\bar{d}(D) = 0$), maka variasi tersebut tidak memengaruhi nilai reliabilitas instrumen. Dengan demikian, seluruh butir soal tetap memenuhi kriteria reliabel berdasarkan metode Percentage of Agreement.

Secara keseluruhan, hasil analisis menunjukkan bahwa instrumen tes keterampilan berpikir tingkat tinggi pada materi kinematika memiliki reliabilitas yang sangat tinggi. Hal ini menunjukkan bahwa instrumen memiliki tingkat konsistensi penilaian antarvalidator yang sangat baik sehingga layak digunakan pada tahap uji coba empiris kepada peserta didik. Tingginya nilai reliabilitas juga mengindikasikan bahwa instrumen mampu memberikan hasil penilaian yang stabil dan konsisten apabila digunakan oleh validator yang berbeda.

Berdasarkan hasil tersebut, dapat disimpulkan bahwa seluruh 20 butir soal dinyatakan reliabel dan memenuhi persyaratan sebagai instrumen yang konsisten untuk mengukur keterampilan berpikir tingkat tinggi pada materi kinematika. Oleh karena itu, instrumen dapat digunakan pada tahap pengujian selanjutnya, yaitu analisis validitas empiris, tingkat kesukaran, daya pembeda, efektivitas pengecoh, dan analisis karakteristik butir soal terhadap peserta didik.

Tabel 2. Hasil Analisis Reliabilitas Tes KBK Tiap Soal

No. Soal	$\overline{d(A)}$	$\overline{d(D)}$	Nilai R(%)	Kesimpulan
1	3	0	10	Reliabel
2	3	0	100	Reliabel
3	4	0	100	Reliabel
4	3	0	10	Reliabel
5	4	0	100	Reliabel
6	3,5	0	100	Reliabel
7	4	0	100	Reliabel
8	4	0	100	Reliabel
9	3,5	0	100	Reliabel
10	3,5	0	100	Reliabel
11	3	0	100	Reliabel
12	4	0	100	Reliabel
13	3,5	0	100	Reliabel
14	3,5	0	100	Reliabel
15	4	0	100	Reliabel
16	4	0	100	Reliabel
17	3,5	0	100	Reliabel
18	3	0	100	Reliabel
19	3	0	100	Reliabel
20	3,5	0	100	Reliabel

Tingkat Kesukaran Tes KBK

Hasil analisis tingkat kesukaran dari 20 butir soal disajikan dalam Tabel 3.

Tabel 3. Distribusi Tingkat Kesukaran Soal KBK

Persentase(%)	Kategori	Nomor Soal	Jumlah	Persentase (%)	Rata-Rata (%)
27	Mudah	1, 2	2	10	20, 5
28-72	Sedang	3, 4, 5, 6, 7,8, 9, 10, 11, 13, 14, 16, 17, 18, 19, 20	16	80	52, 87
≥ 73	Sukar	12, 15	2	10	87, 5

Hasil Analisis Tingkat Kesukaran Butir Soal Tes Keterampilan Berpikir Kreatif (KBK)

Tingkat kesukaran merupakan salah satu karakteristik penting dalam analisis butir soal yang digunakan untuk mengetahui tingkat kemudahan atau kesulitan setiap butir soal bagi peserta didik. Analisis tingkat kesukaran dilakukan terhadap 20 butir soal tes KBK pada materi kinematika. Hasil analisis disajikan pada Tabel 3.

Berdasarkan Tabel 3, diketahui bahwa terdapat 2 butir soal (10%) yang termasuk dalam kategori mudah, yaitu soal nomor 1 dan 2, dengan rata-rata indeks kesukaran sebesar 20,5%. Soal-soal pada kategori ini dapat dijawab dengan benar oleh sebagian besar peserta didik sehingga memiliki tingkat kesulitan yang rendah. Kehadiran soal mudah diperlukan untuk mengukur penguasaan konsep dasar serta memberikan kesempatan kepada peserta didik dengan kemampuan rendah untuk memperoleh skor.

Sebanyak 16 butir soal (80%) berada pada kategori sedang, yaitu soal nomor 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 16, 17, 18, 19, dan 20, dengan rata-rata indeks kesukaran sebesar 52,87%. Dominasi butir soal pada kategori sedang menunjukkan bahwa sebagian besar soal memiliki tingkat kesulitan yang proporsional sehingga mampu membedakan kemampuan peserta didik secara lebih optimal. Soal dengan tingkat kesukaran sedang umumnya merupakan karakteristik yang diharapkan dalam penyusunan instrumen evaluasi karena dapat memberikan informasi yang lebih akurat mengenai kemampuan peserta didik.

Sementara itu, terdapat 2 butir soal (10%) yang termasuk dalam kategori sukar, yaitu soal nomor 12 dan 15, dengan rata-rata indeks kesukaran sebesar 87,5%. Soal-soal tersebut hanya dapat dijawab dengan benar oleh sebagian kecil peserta didik sehingga memiliki tingkat kesulitan yang relatif tinggi. Meskipun demikian, keberadaan soal sukar tetap diperlukan dalam suatu instrumen untuk mengidentifikasi peserta didik yang memiliki kemampuan berpikir tingkat tinggi dan penguasaan materi yang lebih baik.

Secara keseluruhan, distribusi tingkat kesukaran menunjukkan bahwa instrumen didominasi oleh soal berkategori sedang (80%), sedangkan soal berkategori mudah dan sukar masing-masing berjumlah 10%. Komposisi ini menunjukkan bahwa tes telah memiliki sebaran tingkat kesukaran yang cukup seimbang dengan proporsi terbesar berada pada kategori sedang. Kondisi tersebut mengindikasikan bahwa instrumen telah memenuhi salah satu karakteristik tes yang baik, yaitu memiliki mayoritas butir soal dengan tingkat kesukaran sedang sehingga mampu mengukur kemampuan peserta didik secara efektif dan menghasilkan informasi yang lebih representatif mengenai keterampilan berpikir tingkat tinggi pada materi kinematika.

Daya Pembeda Tes KBK

Tabel 4. Daya Pembeda Soal KBT

No	Daya Pembeda	No. Soal	Jumlah	Persentase (%)	Rata-Rata
1	ID>0.40 (Sangat Baik)	3,4,5,6,7,8,9,10,20	11	55	0.50
2	0.30<ID<0.39 (Cukup Baik)	2,17	2	10	0.41
3	0.20<ID<0.29 (Kurang Baik)	13,18	1	5	0.41
4	ID< 0.20 (Tidak Baik)	1,12,14,15,16,19	6	30	0.07

Hasil Analisis Daya Pembeda Butir Soal Tes KBK

Daya pembeda merupakan kemampuan suatu butir soal untuk membedakan peserta didik yang memiliki kemampuan tinggi dengan peserta didik yang memiliki kemampuan rendah. Semakin tinggi indeks daya pembeda suatu butir soal, semakin baik kemampuan soal tersebut dalam mengidentifikasi perbedaan kemampuan peserta didik. Hasil analisis daya pembeda terhadap 20 butir soal tes keterampilan berpikir tingkat tinggi (KBT) pada materi kinematika disajikan pada Tabel 4.

Berdasarkan Tabel 4, diperoleh bahwa sebanyak 11 butir soal (55%) memiliki daya pembeda sangat baik ($ID > 0,40$), yaitu soal nomor 3, 4, 5, 6, 7, 8, 9, 10, 11, dan 20, dengan rata-rata indeks daya pembeda sebesar 0,50. Hasil ini menunjukkan bahwa lebih dari separuh butir soal mampu membedakan secara optimal antara peserta didik yang berkemampuan tinggi dan peserta didik yang berkemampuan rendah. Soal-soal pada kategori ini layak dipertahankan karena memiliki kualitas yang sangat baik sebagai alat evaluasi.

Selanjutnya, terdapat 2 butir soal (10%) yang berada pada kategori cukup baik ($0,30 \leq ID < 0,40$), yaitu soal nomor 2 dan 17, dengan rata-rata indeks daya pembeda sebesar 0,41. Meskipun berada di bawah kategori sangat baik, butir-butir soal tersebut masih mampu membedakan kemampuan peserta didik dengan cukup baik sehingga tetap layak digunakan, meskipun dapat dipertimbangkan untuk dilakukan penyempurnaan.

Pada kategori kurang baik ($0,20 \leq ID < 0,30$), terdapat 2 butir soal, yaitu nomor 13 dan 18. Butir soal pada kategori ini memiliki kemampuan yang relatif rendah dalam membedakan peserta didik berdasarkan tingkat kemampuannya. Oleh karena itu, soal-soal tersebut sebaiknya direvisi agar memiliki karakteristik pembeda yang lebih baik sebelum digunakan kembali dalam pengukuran.

Sementara itu, terdapat 6 butir soal (30%) yang termasuk kategori tidak baik ($ID < 0,20$), yaitu soal nomor 1, 12, 14, 15, 16, dan 19, dengan rata-rata indeks daya pembeda sebesar 0,07. Nilai tersebut menunjukkan bahwa butir soal belum mampu membedakan peserta didik yang berkemampuan tinggi dan rendah secara efektif. Rendahnya daya pembeda dapat disebabkan oleh beberapa faktor, seperti tingkat kesukaran yang tidak proporsional, adanya alternatif jawaban yang kurang berfungsi, atau rumusan soal yang kurang jelas. Oleh karena itu, butir-butir soal pada kategori ini perlu direvisi secara menyeluruh atau diganti dengan butir soal baru.

Secara keseluruhan, hasil analisis menunjukkan bahwa sebagian besar butir soal (55%) telah memiliki daya pembeda yang sangat baik. Namun, masih terdapat beberapa butir soal yang berada pada kategori kurang baik dan tidak baik sehingga memerlukan perbaikan sebelum dijadikan sebagai bagian dari instrumen tes yang terstandarisasi. Dengan demikian, hasil analisis daya pembeda memberikan informasi penting dalam proses penyempurnaan instrumen sehingga setiap butir soal yang digunakan benar-benar mampu mengukur keterampilan berpikir tingkat tinggi peserta didik secara efektif.

Hasil penelitian menunjukkan bahwa instrumen tes keterampilan berpikir tingkat tinggi (KBT) pada materi kinematika memiliki kualitas yang baik ditinjau dari aspek validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Keempat aspek tersebut merupakan karakteristik psikometrik yang penting dalam proses standarisasi instrumen sehingga instrumen yang dihasilkan mampu memberikan informasi yang akurat mengenai kemampuan peserta didik (Zain et al., 2022).

Hasil validasi menggunakan koefisien konsistensi internal Gregory menunjukkan bahwa validitas konstruk dan validitas isi memperoleh koefisien sebesar 1,00 dengan kategori sangat tinggi. Nilai tersebut menunjukkan adanya kesepakatan sempurna antarvalidator mengenai kesesuaian setiap butir soal dengan indikator keterampilan berpikir tingkat tinggi, materi kinematika, serta kaidah penyusunan instrumen evaluasi. Semakin tinggi koefisien Gregory, semakin tinggi pula tingkat kesepakatan para ahli terhadap relevansi butir soal, sehingga instrumen dapat dinyatakan memiliki validitas isi yang sangat baik. Koefisien Gregory pada rentang 0,80–1,00 dikategorikan sebagai validitas isi sangat tinggi.

Temuan tersebut sejalan dengan penelitian Kurniawan, Sinaga, dan Saepuzaman yang menyatakan bahwa instrumen keterampilan berpikir kreatif yang memiliki validitas isi dan konstruk tinggi lebih mampu merepresentasikan indikator kemampuan yang akan diukur sehingga layak digunakan dalam proses evaluasi pembelajaran. Validitas yang tinggi menunjukkan bahwa setiap butir soal telah mengukur konstruk yang sesuai dan mampu merepresentasikan kompetensi berpikir tingkat tinggi secara tepat (Pada et al., 2018).

Selain validitas, hasil analisis reliabilitas menggunakan Percentage of Agreement (PoA) menunjukkan bahwa seluruh butir soal memperoleh nilai reliabilitas sebesar 100%. Nilai tersebut menunjukkan bahwa seluruh validator memberikan penilaian yang konsisten terhadap setiap butir soal tanpa adanya perbedaan pendapat (disagreement = 0). Tingginya nilai reliabilitas menunjukkan bahwa instrumen memiliki konsistensi yang sangat baik sehingga apabila digunakan oleh validator yang berbeda akan memberikan hasil penilaian yang relatif sama. Dalam metode Percentage of Agreement, tingkat kesepakatan di atas 75% menunjukkan bahwa instrumen telah memenuhi kriteria reliabel (Arifin, 2012; Zainuri et al., 2022).

Berdasarkan analisis tingkat kesukaran, sebagian besar butir soal (80%) berada pada kategori sedang, sedangkan masing-masing 10% termasuk kategori mudah dan sukar. Dominasi soal berkategori sedang menunjukkan bahwa instrumen telah memiliki tingkat kesulitan yang proporsional. Butir soal dengan tingkat kesukaran sedang umumnya mampu memberikan informasi yang lebih baik mengenai kemampuan peserta didik karena tidak terlalu mudah maupun terlalu sulit. Kondisi ini menunjukkan bahwa sebagian besar soal telah sesuai untuk mengukur keterampilan berpikir tingkat tinggi pada materi kinematika. Hasil penelitian ini

juga sejalan dengan penelitian pengembangan instrumen kemampuan berpikir kreatif fisika yang melaporkan bahwa instrumen berkualitas umumnya didominasi oleh butir soal dengan tingkat kesukaran sedang sehingga mampu memberikan hasil pengukuran yang lebih akurat.

Hasil analisis daya pembeda menunjukkan bahwa 55% butir soal berada pada kategori sangat baik, 10% berkategori cukup baik, sedangkan sebagian butir lainnya masih berada pada kategori kurang baik dan tidak baik. Dominasi butir soal dengan daya pembeda sangat baik menunjukkan bahwa sebagian besar soal mampu membedakan peserta didik yang memiliki kemampuan tinggi dengan peserta didik yang memiliki kemampuan rendah. Hal ini mengindikasikan bahwa instrumen telah memiliki kemampuan diskriminatif yang baik sehingga layak digunakan sebagai alat ukur keterampilan berpikir tingkat tinggi. Sebaliknya, butir soal yang memiliki daya pembeda rendah perlu direvisi karena belum mampu mengidentifikasi perbedaan kemampuan peserta didik secara optimal.

Keberadaan beberapa butir soal dengan daya pembeda rendah dapat disebabkan oleh beberapa faktor, seperti tingkat kesukaran yang terlalu tinggi atau terlalu rendah, alternatif jawaban yang kurang berfungsi, ataupun redaksi soal yang masih menimbulkan penafsiran ganda. Oleh karena itu, revisi terhadap butir-butir tersebut diperlukan agar kualitas instrumen semakin baik. Proses revisi merupakan bagian penting dalam standarisasi instrumen karena bertujuan menghasilkan perangkat tes yang memenuhi persyaratan psikometrik sebelum digunakan secara luas (Tawil & Liliasari, 2025)

Secara keseluruhan, hasil penelitian menunjukkan bahwa instrumen tes keterampilan berpikir tingkat tinggi pada materi kinematika telah memenuhi sebagian besar karakteristik instrumen yang baik, yaitu memiliki validitas isi dan konstruk yang sangat tinggi, reliabilitas yang sangat tinggi, dominasi tingkat kesukaran pada kategori sedang, serta daya pembeda yang sebagian besar berada pada kategori sangat baik. Hasil tersebut menunjukkan bahwa instrumen layak digunakan untuk mengukur keterampilan berpikir tingkat tinggi peserta didik. Meskipun demikian, beberapa butir soal yang memiliki daya pembeda rendah masih memerlukan penyempurnaan agar kualitas instrumen semakin optimal dan memenuhi standar instrumen evaluasi pendidikan.

Ringkasan Substansi Hasil Penelitian

Hasil penelitian menunjukkan bahwa produk tes KBK yang dikembangkan memenuhi syarat kualitas produk sesuai teori Neeven, yakni kevalidan, reliabilitas isi dan konstruk, Tingkat Kesukaran, dan daya pembeda. Hal itu menunjukkan bahwa semua komponen-komponennya saling berkaitan, teori-teori belajar dan pembelajaran, dan teori standarisasi tes yang dikembangkan.

Analisis Hubungan dengan Penelitian Lain

Temuan penelitian ini memperlihatkan keterkaitan kuat dengan berbagai studi terdahulu mengenai asesmen berpikir kreatif, termasuk salah satunya adalah literasi. Misalnya, jika dibandingkan dengan hasil penelitian yang dilakukan oleh Lestari., Supardi., & Jatmiko, (2022) yang menyoroti bahwa efektivitas asesmen berpikir kreatif sangat menentukan keberhasilan peserta didik memahami konsep-konsep dan memecahkan berbagai masalah yang dihadapinya. Demikian pula temuan (Nur., Dewanti, & Santoso, 2024; Li et al, 2024; Qiao, et al, 2024; Matraeva et al., 2020) mengkaji kreativitas seseorang sangat ditentukan oleh keterampilan problem solving.

Refleksi Manfaat Tujuan Penelitian

Refleksi terhadap hasil penelitian menunjukkan bahwa standarisasi tes KBK memiliki manfaat besar bagi perguruan tinggi, dosen, guru fisika dan peserta didik SMN kelas IX dalam mengukur salah satu variabel berpikir tingkat tinggi. Hasil pengukuran tes KBK ini akan lebih tepat dan sesuai potensi yang dimiliki oleh peserta didik, sehingga mereka dapat mengembangkan dirinya berkekrativitas pada proses pembelajaran di kelas dan juga bermanfaat untuk memberi kesempatan para peserta didik menentukan pilihan kerja di masa akan datang. Hal ini sesuai dengan temuan penelitian bahwa dengan menguasai indikator-indikator KBK, maka akan lebih mudah mengembangkan daya kreativitas peserta didik (Ramadhan et al, 2026; Reski et al, 2024; Matraeva et al., 2020).

Analisis Implikasi Hasil Penelitian

Implikasi dari hasil penelitian ini mengarah pada pentingnya pengembangan KBK yang berkualitas, dalam melakukan evaluasi KBK seseorang. Hasil ini dapat dijadikan pedoman untuk merancang kurikulum yang mengembangkan berpikir tingkat tinggi, terutama untuk mencapai delapan dimensi kelulusan saat sekarang ini (OECD, 2021., Suyanto., et al, 2025; Son, M., & Ha, .(2025).

Analisis Penyebab Fenomena Hasil Penelitian

Fenomena hasil penelitian yang menunjukkan keberhasilan peserta didik mengembangkan KBK mereka, untuk mencapai hal tersebut, maka penting dilatihkan semua indikator-indikatornya secara kontinu dan berkelanjutan. Kegiatan latihan tersebut dikerjakan pada lembar kegiatan peserta didik (LKPD). Pada proses ini, mereka secara nyata dapat melihat keterkaitan antara fenomena IPA dengan berbagai jenis variabel, Seperti halnya tuuan yang menegaskan pentingnya sinergi antara KBK dengan kegiatan saintifik. (Abu et al., 2025; Baihaqi et al, 2025; Lubis., Pantiwati., & Rahardjanto, 2025; Haliza., Haka., & Pratama, 2025).

Analisis Aksi Berdasarkan Hasil Penelitian

Berdasarkan keseluruhan temuan, beberapa tindakan strategis perlu diambil untuk memperkuat keberhasilan peserta didik meningkatkan KBK di SMAN. Pertama, sekolah perlu memperluas pelatihan pembuat soal KBK, bagi guru fisika dan dosen IPA (fisika, kimia, dan biologi) pada umumnya agar mereka memiliki keterampilan membuat tes yang tersandarisasi. Kedua, perlu adanya penguatan sistem kurikulum IPA yang menekankan pengembangan berpikir tingkat tinggi, khususnya KBK dan penerapan pada proses pembelajaran di kelas secara kontinu dan komprehensif.

SIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa instrumen tes keterampilan berpikir kreatif (KBK) pada materi kinematika memiliki kualitas yang baik sebagai instrumen evaluasi pembelajaran. Hasil validasi menggunakan koefisien Gregory menunjukkan bahwa validitas konstruk dan validitas isi masing-masing memperoleh koefisien 1,00 dengan kategori sangat tinggi, sehingga setiap butir soal dinyatakan telah sesuai dengan indikator KBK, materi kinematika, serta kaidah penyusunan instrumen. Hasil analisis reliabilitas menggunakan Percentage of Agreement (PoA) menunjukkan bahwa seluruh butir soal memperoleh nilai reliabilitas 100%, yang mengindikasikan adanya tingkat kesepakatan sempurna antarvalidator. Dengan demikian, instrumen memiliki konsistensi yang sangat tinggi dan layak digunakan sebagai alat ukur. Analisis tingkat kesukaran menunjukkan bahwa 80% butir soal berada pada kategori sedang, sedangkan masing-masing 10% berada pada kategori mudah dan sukar. Distribusi tersebut menunjukkan bahwa instrumen memiliki komposisi tingkat kesukaran yang proporsional sehingga mampu mengukur kemampuan peserta didik secara efektif. Berdasarkan analisis daya pembeda, sebagian besar butir soal (55%) memiliki daya pembeda sangat baik, sedangkan sebagian lainnya berada pada kategori cukup baik, kurang baik, dan tidak baik. Hasil ini menunjukkan bahwa sebagian besar soal telah mampu membedakan peserta didik yang memiliki kemampuan tinggi dan rendah, meskipun beberapa butir soal masih memerlukan revisi agar kualitas instrumen semakin optimal. Secara keseluruhan, instrumen tes KBK pada materi kinematika telah memenuhi sebagian besar karakteristik instrumen yang baik, yaitu valid, reliabel, memiliki tingkat kesukaran yang didominasi kategori sedang, serta daya pembeda yang sebagian besar sangat baik. Oleh karena itu, instrumen ini layak digunakan sebagai salah satu alat evaluasi untuk mengukur keterampilan berpikir tingkat tinggi peserta didik pada materi kinematika setelah dilakukan penyempurnaan terhadap beberapa butir soal yang masih memiliki daya pembeda rendah.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada guru fisika SMN di Kabuapten Gowa, dan juga kepada seluruh responden yang telah memberikan jawaban soal-soal dengan penuh kepercayaan, dan tanggungjawabnya. Demikian pula pakar yang telah melakukan validasi terhadap instrumen penelitian ini. kerjasamanya sehingga kegiatan ini dapat terlaksana.

DAFTAR PUSTAKA

- Arifin, Z. (2012). *Evaluasi Pembelajaran*. PT. Remaja Rosdakarya.
- Azman, N. A., Hamzah, M. I., & Baharudin, H. (2025). Digital Teaching Strategies of Islamic Education Teachers: A Case Study in Primary Schools. *International Journal of Learning, Teaching and Educational Research*. <https://doi.org/10.26803/ijlter.24.3.27>
- Gregory, R. J. (2004). *Psychological Testing: History, Principles, and Applications* (4th ed.). Pearson.
- Kurniawan, K., Sinaga, P., & Saepuzaman, D. (2025). Aiken's V analysis and Rasch modeling to determine the quality of physics creative thinking test instruments. *Jurnal Riset Dan Kajian Pendidikan Fisika*, 12(1). <https://doi.org/10.12928/jrkpf.v12i1.1161>
- Noviani, J., Ahmad, N. Q., Dewi, I. P., & Ulfa, M. (2025). Exploring the influence of Sternberg's thinking styles on students' mathematical creative problem-solving. *Alifmatika*.
- OECD. (2024). PISA 2022 Results (Volume III): Creative Minds, Creative Schools. In PISA. OECD Publishing. <https://doi.org/10.1787/765ee8c2-en>
- Pada, A. U. T., Mustakim, S. S., & Subali, B. (2018). Construct validity of creative thinking skills instrument for biology student teachers in the subject of human physiology. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 22(2), 119–129. <https://doi.org/10.21831/pep.v22i2.22369>
- Pradana. (2025). Penggunaan model Gregory untuk validitas isi instrumen pembelajaran matematika. *Indonesian Journal of Education & Mathematical Science*.
- Ramdani, A., Artayasa, I. P., Yustiqvar, M., & Nisrina, N. (2021). Enhancing Prospective Teachers' Creative Thinking Skills: a Study of the Transition From Structured To Open Inquiry Classes. *Cakrawala Pendidikan*, 40(3), 637–649. <https://doi.org/10.21831/cp.v40i3.41758>
- Santofani, A., & Rosana, D. (2016). Pengembangan tes kreativitas pada pembelajaran fisika dengan pendekatan inkuiri pada materi teori kinetik gas. *Jurnal Inovasi Pendidikan IPA*, 2(2), 134–144. <https://doi.org/10.21831/jipi.v2i2.6373>
- Tanjung, Y. I., & Nasution, I. R. (2022). The Development of Creative Thinking Test Instruments with Torrance Indicators on Direct Current Electricity Materials. *Jurnal Pendidikan Fisika Indonesia*, 18(2). <https://doi.org/10.15294/jpfi.v18i2.32117>
- Tawil, M., & Liliarsari. (2025). *Teori dan Implementasi Pembelajaran IPA*. Badan Penerbit UNM.
- Tiadarma, M. A., Sahyar, & Bunawan, W. (2022). Development of creative thinking skills test instruments on kinematics materials for SMA/MA students in Medan. *AIP Conference Proceedings*, 2659, 120006. <https://doi.org/10.1063/5.0113166>
- Utama, A. S., Rosidin, U., & Suyatna, A. (2021). Instrument of Creative Thinking Skills in Project Based Learning Physics. *Indonesian Journal of Science and Mathematics Education*, 4(1), 58–68. <https://doi.org/10.24042/ijsme.v4i1.8572>
- Zain, A. Z., Ramalis, T. R., & Muslim. (2022). Karakterisasi Instrumen Tes Keterampilan Berpikir Kreatif Berdasarkan Analisis Partial Credit Model. *Jurnal Ilmiah Pendidikan Fisika*, 6(1), 176–187. <https://doi.org/10.20527/jipf.v6i1.4806>

1071 *Standarisasi Tes Keterampilan Berpikir Kreatif Kinematika - Muh. Tawil, Haksan Darwangsa*
DOI : <https://doi.org/10.31004/edukatif.v8i5.9412>

Zainuri, A., Aquami, & Annur, S. (2022). *Evaluasi Pendidikan (Kajian Teoritik)*. Qiara Media.