

Desain Instrumen Tes Pemahaman Konsep Mendalam pada Topik Dinamika Partikel Satu Dimensi: Validasi Isi dan Empiris

Muhammad Reyza Arief Taqwa¹, Parlindungan Sinaga¹✉,
Endi Suhendi¹, Selly Feranie¹, Handy Faishal Rahim²

¹Universitas Pendidikan Indonesia

Jl. Dr. Setiabudi No.229, Kota Bandung, Jawa Barat 40154, Indonesia

²Universitas Negeri Malang

Jl. Semarang No.5, Sumbersari, Kota Malang, Jawa Timur 65145, Indonesia

| psubaga@upi.edu ✉ | DOI : <https://doi.org/10.37729/jips.v7i1.7502> |

Article Info

Submitted

11/03/2026

Revised

27/04/2026

Accepted

15/05/2026

Abstrak – Pemahaman konsep mendalam pada topik dinamika partikel satu dimensi menjadi fondasi esensial pada mata kuliah Mekanika Klasik. Disisi lain, belum ada instrumen yang dapat diadaptasi secara luas untuk mengukur kemampuan tersebut. Penelitian ini bertujuan untuk mendesain dan memvalidasi instrumen tes pemahaman konsep mendalam pada topik dinamika partikel satu dimensi. Pendekatan kuantitatif digunakan dengan desain non-eksperimental dan *cross-sectional*, melalui tiga tahap utama, yaitu validasi ahli, revisi instrumen, dan uji empiris pada mahasiswa fisika dan pendidikan fisika jenjang sarjana. Validitas isi dievaluasi menggunakan *Content Validity Index* (CVI), sedangkan validitas empiris butir dianalisis melalui korelasi *point-biserial* (r_{pbis}) dan reliabilitas instrumen diuji menggunakan Bayesian unidimensional reliability. Hasil penelitian menunjukkan bahwa seluruh butir memiliki validitas isi yang sangat baik, dengan dua butir mengalami revisi berdasarkan masukan ahli agar lebih tepat secara konseptual dan representasional. Uji empiris menunjukkan bahwa seluruh butir soal valid, dengan nilai r_{pbis} yang signifikan dan berada pada kategori daya beda yang baik, serta instrumen memiliki konsistensi internal yang baik ($\alpha = 0,789$; 95% CI: 0,758–0,841) berdasarkan estimasi reliabilitas Bayesian. Temuan ini mengimplikasikan bahwa instrumen yang dikembangkan layak digunakan untuk menilai pemahaman konsep mendalam mahasiswa pada topik dinamika partikel satu dimensi, sekaligus berpotensi menjadi acuan dalam evaluasi pembelajaran fisika dan pengembangan instrumen serupa pada topik lain.

Kata kunci: Instrumen tes, Pemahaman konsep mendalam, Dinamika partikel satu dimensi

Abstract - A deep conceptual understanding of one-dimensional particle dynamics is an essential foundation for the Classical Mechanics course. However, there is currently no widely adaptable instrument available to measure this ability. This study aims to design and validate a test instrument for deep conceptual understanding of one-dimensional particle dynamics. A quantitative approach was used with a non-experimental, cross-sectional design, involving three main stages: expert validation, instrument revision, and empirical testing among undergraduate physics and physics education students. Content validity was evaluated using the Content Validity Index (CVI), while the empirical validity of the items was analyzed via point-biserial correlation (r_{pbis}), and the instrument's reliability was tested using Bayesian unidimensional reliability. The results showed that all items had excellent content validity, with two items undergoing revision based on expert input to ensure greater conceptual and representational accuracy. Empirical testing indicated that all test items were valid, with significant r_{pbis} values falling within the "good discriminant power" category, and the instrument demonstrated good internal consistency ($\alpha = 0.789$; 95% CI: 0.758–0.841) based on Bayesian reliability estimates. These findings suggest that the instrument developed is suitable for assessing students' deep conceptual understanding of one-dimensional particle dynamics, and has the potential to serve as a reference for evaluating physics instruction and developing similar instruments for other topics.



Keywords: Test instruments, Deep conceptual understanding, One-dimensional particle dynamics

1. Pendahuluan

Pemahaman konsep yang mendalam dalam topik dinamika partikel satu dimensi merupakan fondasi esensial dalam perkuliahan Mekanika Klasik di tingkat perguruan tinggi. Berbeda dengan pengantar fisika dasar yang lebih menekankan pada konsep-konsep fundamental dan aplikasi sederhana, mekanika klasik menuntut mahasiswa untuk mampu menganalisis gerak benda secara matematis dan abstrak [1], yang selanjutnya dituntut untuk mampu menghubungkan hukum-hukum Newton dengan konsep lain seperti momentum, energi, serta formulasi Lagrangian atau Hamiltonian yang lebih kompleks. Dinamika partikel satu dimensi, sebagai fondasi awal perkuliahan, menjadi gerbang bagi mahasiswa untuk memahami sistem fisis yang lebih kompleks dan rumit [2]–[4]. Ketidakmampuan menguasai topik ini secara mendalam akan berdampak signifikan terhadap pemahaman materi selanjutnya seperti osilasi, gelombang, atau mekanika dalam tiga dimensi [5]–[7]. Lebih lanjut, penelitian di bidang pendidikan fisika secara konsisten menunjukkan bahwa kesulitan mahasiswa pada topik-topik fondasi mekanika bersifat persisten dan cenderung terbawa ke dalam pemahaman topik lanjutan [8]. Oleh karena itu, diperlukan adanya instrumen yang mampu memetakan sejauh mana mahasiswa telah mencapai pemahaman konsep yang mendalam, bukan sekadar hafalan rumus atau prosedur penyelesaian soal.

Karakteristik utama dari pemahaman konsep mendalam adalah kemampuan seseorang untuk merepresentasikan suatu konsep dalam berbagai modus representasi serta keluwesan dalam mentranslasikannya dari satu modus ke modus lainnya. Translasi ini menjadi keterampilan penting karena merupakan proses berpikir yang mengubah suatu konsep dari satu bentuk representasi ke bentuk representasi lain yang berbeda, tanpa mengubah makna dari konsep tersebut [9]. Hasil penelitian terdahulu menemukan bahwa banyak mahasiswa dapat memahami suatu konsep dalam representasi tertentu, namun gagal ketika dinyatakan dalam representasi yang berbeda [10], [11]. Dalam konteks dinamika partikel satu dimensi, sebuah konsep dapat direpresentasikan melalui diagram gerak (*pictorial*), grafik, persamaan matematis (representasi simbolik), tabel, atau bahkan deskripsi verbal. Mahasiswa yang memiliki pemahaman mendalam seharusnya tidak hanya mampu membaca sebuah grafik kecepatan terhadap waktu, tetapi juga mampu mentranslasikannya ke dalam persamaan gerak, mendeskripsikan gaya-gaya yang bekerja dalam bentuk diagram benda bebas, serta menceritakan kembali fenomena fisis yang terjadi dalam bentuk narasi yang koheren. Kemampuan translasi antarrepresentasi inilah yang menjadi indikator kunci bahwa mahasiswa benar-benar memahami esensi dari konsep dinamika secara mendalam, bukan sekadar terampil dalam satu jenis soal tertentu.

Urgensi pemahaman konsep secara mendalam, khususnya dalam bidang pendidikan fisika, telah lama disadari. Namun demikian, penelitian dan pengembangan instrumen tes yang secara eksplisit dirancang untuk mengukur kedalaman pemahaman konsep pada topik dinamika partikel, khususnya di jenjang perkuliahan lanjut (mekanika klasik), masih sangat terbatas. Selama ini, pengembangan instrumen lebih banyak difokuskan pada ranah Fisika Dasar atau pengantar, seperti *Force Concept Inventory* (FCI) [12]. Instrumen-instrumen yang ada mayoritas dirancang untuk mengidentifikasi miskonsepsi umum pada tingkat pengantar (misalnya, Dewi dkk. [13] dan Ernanda dkk. [14]), belum secara spesifik menyentuh kedalaman analisis yang dibutuhkan dalam perkuliahan mekanika klasik. Instrumen yang ada selama ini, seperti *three-tier diagnostic test* dan instrumen diagnostik konsepsi [15], [16], dirancang terutama untuk mengidentifikasi miskonsepsi pada tingkat pengetahuan deklaratif atau aplikasi langsung, bukan untuk mengukur kemampuan translasi antarrepresentasi yang menjadi inti dari pemahaman konsep mendalam. Dengan demikian, celah penelitian yang spesifik teridentifikasi: belum tersedia instrumen yang secara konstruksi mengintegrasikan multiple external representations (MERs) dan menuntut kemampuan translasi antarrepresentasi sebagai indikator utama pemahaman konsep mendalam pada topik dinamika partikel satu dimensi di jenjang perkuliahan mekanika klasik.

Penelitian ini hadir untuk mengisi celah tersebut dengan mendesain dan memvalidasi instrumen yang secara sengaja mengoperasionalkan kedalaman pemahaman melalui tuntutan translasi antarmodus representasi, yakni dari representasi grafik, persamaan matematis, diagram gerak, tabel, hingga deskripsi verbal.

Oleh karena itu, penelitian ini bertujuan untuk mendesain dan memvalidasi instrumen tes pemahaman konsep mendalam pada topik dinamika partikel 1 dimensi untuk mata kuliah mekanika klasik. Instrumen ini dirancang secara khusus dengan menyajikan soal-soal dalam berbagai modus representasi yang menuntut mahasiswa untuk melakukan translasi antar representasi, sehingga diharapkan mampu mengukur kemampuan kognitif tingkat tinggi mahasiswa secara komprehensif.

2. Metode

2.1. Desain dan Prosedur Penelitian

Penelitian ini dilakukan dalam kerangka paradigma positivisme, dengan menggunakan metodologi kuantitatif, desain non-eksperimental, dan pendekatan *cross-sectional* dengan cakupan deskriptif. Prosedur penelitian terdiri atas tiga tahap: pertama, draf instrumen divalidasi oleh para ahli pendidikan fisika; kedua, dilakukan revisi terhadap draf instrumen berdasarkan saran perbaikan dari para ahli; dan ketiga, instrumen yang telah direvisi diujicobakan secara empiris kepada mahasiswa.

2.2. Sumber Data

Sumber data utama dalam penelitian ini adalah ahli (*expert*) dalam bidang pendidikan fisika dan mahasiswa fisika dan pendidikan fisika jenjang sarjana. Ahli berperan untuk menilai validitas konstruk dan isi dari instrumen yang telah dirancang. Pelibatan ahli dari empat perguruan tinggi berbeda bertujuan untuk meningkatkan keragaman perspektif dan memperluas sudut pandang dalam penilaian instrumen. Di sisi lain, mahasiswa menjadi subjek uji coba untuk memperoleh data empiris terkait kualitas butir soal. Dalam penelitian ini, mahasiswa yang menjadi sumber data merupakan mahasiswa yang sedang atau telah selesai menempuh mata kuliah mekanika klasik. Bagi mahasiswa yang sedang proses menempuh mata kuliah mekanika klasik dipastikan telah selesai membahas topik dinamika partikel satu dimensi. Mahasiswa yang menjadi sumber data pada uji empiris ini berasal dari empat universitas, dua diantaranya dari universitas di pulau Jawa, sedangkan dua lainnya dari universitas di pulau Sumatera dan Kalimantan. Sebaran sumber data ini diharapkan dapat memperluas generalisasi temuan karena didasarkan dari sumber data yang heterogen. Informasi lebih lengkap mengenai sumber data ditunjukkan oleh Tabel 1.

Tabel 1. Informasi Sumber Data

Aspek	Sub-aspek	<i>f</i>	%
A. Ahli (<i>expert</i>) (<i>n</i> = 5)			
Jenjang Pendidikan	Doktoral	5	100,00%
Masa kerja (tahun)	0-5	0	0,00%
	5-10	1	20,00%
	> 10	4	80,00%
B. Mahasiswa (<i>n</i> = 189)			
Jenis kelamin	Laki-laki	32	16,93%
	Perempuan	157	83,07%
Semester	II	86	45,50%
	IV	67	35,45%
	VI	36	19,05%
Program Studi	S1 Pendidikan Fisika	102	53,97%
	S1 Fisika	87	46,03%

2.3. Rancangan Instrumen

Validasi instrumen oleh validator menggunakan skala likert dengan skor 1 (tidak sesuai), 2 (kurang sesuai), 3 (cukup sesuai), dan 4 (sesuai) untuk menilai validitas instrumen pemahaman konsep mendalam. Dalam panduan validasi diberikan informasi panduan pemberian skor yakni, skor 1 jika item tidak sesuai seluruhnya berdasarkan pernyataan sehingga harus diganti atau seluruhnya direvisi, skor 2 jika item dominan tidak sesuai berdasarkan butir pernyataan sehingga sebagian besar harus direvisi, skor 3 jika item dominan sesuai berdasarkan butir pernyataan namun terdapat beberapa hal yang tidak sesuai sehingga sebagian kecil harus direvisi, dan skor 4 jika item sepenuhnya sesuai berdasarkan butir pernyataan sehingga sebagian dapat digunakan tanpa ada revisi. Adapun aspek yang dinilai adalah kesesuaian konteks pertanyaan, konstruksi, dan bahasa, dengan total sebelas pernyataan. Lembar validasi dapat diakses melalui tautan data pendukung.

Instrumen pemahaman konsep mendalam ini dirancang berdasarkan hasil sintesis dari empat konsep, yakni (1) proses kognitif dalam ranah kognitif pemahaman oleh Anderson & Krathwohl [17], (2) indikator pemahaman konsep oleh Herayanti dkk. [18], (3) pemahaman konsep secara mendalam melalui integrasi beberapa *multiple external representations* (MERs) oleh Ainsworth [19], dan (4) kemampuan menerjemahkan antar representasi untuk membangun abstraksi yang dapat mengarah pada pemahaman yang lebih mendalam terhadap domain yang dipelajari oleh van der Meij & de Jong [20]. Informasi rancangan instrumen tentang nomor soal, aspek berpikir, indikator soal, dan representasi (stimulus soal terhadap opsi jawaban) seperti yang ditunjukkan pada [Apendiks A](#).

2.4. Analisis Data

Dalam penelitian ini dilakukan analisis statistik yang mencakup beberapa prosedur untuk memastikan kualitas psikometrik instrumen yang dirancang. Analisis data dilakukan pada skor penilaian oleh ahli dan skor tes mahasiswa pada uji empiris. Data penilaian validator dianalisis secara kuantitatif dengan menghitung *I-CVI*, *S-CVI/Ave* dan *S-CVI/UA*. Perhitungan dilakukan dengan beberapa tahapan. Pertama, dilakukan perhitungan persetujuan ahli (*expert in agreement*), yakni jumlah ahli yang memberi skor 3 atau 4 [21]. Perhitungan persetujuan ahli dilakukan untuk masing-masing pernyataan di tiap item. Kedua, menentukan skor proporsi *expert agreement* yang merupakan proporsi ahli yang memberi skor 3 atau 4 pada pernyataan di masing-masing item (rasio antara jumlah ahli yang memberi skor 3 atau 4 dengan total ahli). Ketiga, dilakukan perhitungan *item-level content validity index* (*I-CVI*) yang merupakan proporsi ahli yang memberi skor 3 atau 4 untuk tiap item. Dalam penelitian ini, karena setiap item dinilai dengan sebelas pernyataan maka [22] *I-CVI* ditentukan dengan menggunakan persamaan (1).

$$I-CVI = \frac{\text{Jumlah skor proporsi expert agreement}}{\text{Jumlah pernyataan}} \quad (1)$$

Kempat, dilakukan perhitungan *scale-level content validity index based on the average method* (*S-CVI/Ave*) yang merupakan nilai rata-rata skor *I-CVI* untuk seluruh item [23], yang ditentukan dengan menggunakan persamaan (2).

$$S-CVI/Ave = \frac{\Sigma I-CVI}{\text{total pernyataan}} \quad (2)$$

Terakhir, dilakukan perhitungan *scale-level content validity index based on the universal agreement method* (*S-CVI/UA*), yakni proporsi item-item dalam instrumen yang mendapatkan skor 3 atau 4 oleh seluruh ahli secara bersamaan [24]. *Universal Agreement* (UA) merupakan proporsi butir yang mendapat skor 3 atau 4 dari seluruh pakar secara bersamaan, dimana dalam instrumen ini dihitung untuk masing-masing pernyataan di seluruh item. Oleh karena itu, *S-CVI/UA* dihitung dengan menggunakan persamaan (3).

$$S-CVI/UA = \frac{\Sigma UA}{\text{total item} \times \text{total pernyataan}} \quad (3)$$

Data skor tes mahasiswa pada uji empiris dianalisis untuk menentukan validitas dan reliabilitas instrumen. Validitas instrumen ditentukan dengan menggunakan korelasi *point-biserial* karena paling sesuai untuk mengorelasikan variabel dikotomi (skor tiap item) dan variabel interval (total skor tiap item) [25]. Lebih lanjut, menurut Kornbrot [26], koefisien korelasi *point biserial* (r_{pb}) adalah nilai korelasi *Pearson's Product Moment* ketika salah satu variabel bersifat dikotomis sedangkan variabel lainnya bersifat metrik (interval atau rasio). Untuk reliabilitas, analisis dilakukan dengan menggunakan *Bayesian Undimensional Reliability*. Pemilihan analisis ini karena dalam kerangka Bayesian, parameter reliabilitas tidak dianggap sebagai nilai tetap yang tidak diketahui, tetapi sebagai variabel acak yang memiliki distribusi ketidakpastian [27]. Dengan demikian, koefisien reliabilitas yang diperoleh bukan hanya angka tunggal, namun terdapat informasi tambahan distribusi posterior (*posterior distribution*), 95% *credible interval*, dan bahkan probabilitas praktis. Analisis data skor tes mahasiswa ini dilakukan dengan menggunakan aplikasi JASP 0.95.4.0.

3. Hasil dan Pembahasan

3.1. Evaluasi Validitas Instrumen Melalui *Content Validity Index* (CVI)

Hasil perhitungan *I-CVI* berdasarkan skor yang diberikan oleh validator ditunjukkan oleh Tabel 2. Skor *I-CVI* berada dalam interval 0,98-1,00 yang bermakna bahwa keseluruhan item instrumen tes pemahaman konsep mendalam pada topik dinamika partikel 1 dimensi yang telah dikembangkan adalah soal yang valid dengan kategori sangat baik (*excellent*) [23], [28]. Selanjutnya, hasil perhitungan menunjukkan bahwa *S-CVI/Ave* sebesar 0,998 dan *S-CVI/UA* sebesar 0,988 menunjukkan bahwa instrumen secara keseluruhan merupakan instrumen yang valid.

Dalam proses validasi, terdapat dua soal yang menjadi catatan oleh ahli untuk direvisi karena terdapat kesalahan secara konsep, yakni soal nomor 8 dan soal nomor 13. Soal nomor 8 merupakan soal yang menguji proses kognitif membandingkan. Mahasiswa diminta untuk menentukan perbandingan kuadrat kecepatan di tiga posisi berbeda berdasarkan fungsi $F(x)$ (Gambar 1). Sebelum direvisi, soal tersebut menguji kemampuan mahasiswa untuk menentukan posisi benda saat di $x = 0$, $x = 1$ m, dan $x = 4$ m. Secara matematis, soal tersebut dapat diselesaikan, dengan kuadrat kecepatan saat di $x = 0$ dan $x = 4$ m adalah 0, sedangkan kuadrat kecepatan saat di $x = 1$ m adalah $-3 \text{ m}^2/\text{s}^2$ (jawaban benar E).

Tabel 2. Hasil Berpikir Kreatif Tiap Indikator Pada Soal Nomor Tiga

No.Solal	<i>I-CVI</i>	Keterangan
1	1,00	Valid
2	1,00	Valid
3	1,00	Valid
4	1,00	Valid
5	1,00	Valid
6	1,00	Valid
7	1,00	Valid
8	0,98	Valid
9	1,00	Valid
10	1,00	Valid
11	1,00	Valid
12	1,00	Valid
13	0,98	Valid
14	1,00	Valid
15	1,00	Valid

Jika ditinjau secara fisis, kuadrat kecepatan bernilai negatif tidak mungkin terjadi. Artinya, dengan kondisi awal diam dan diberi gaya $F(x) = 2x-4$ maka benda tidak mungkin berada di posisi 1 meter. Kuadrat kecepatan hanya bernilai positif pada posisi $x \leq 0$ atau $x \geq 4$ m. Benda akan bergerak ke kiri jika posisi awal 0 dan benda bergerak ke kanan jika posisi awal 4 m, namun benda tidak mungkin berada pada daerah $0 < x < 4$ m. Berdasarkan saran ahli, soal sebaiknya diubah pertanyaannya untuk x yang mungkin dilalui benda secara fisis. Jika pertanyaan tidak diubah dan hanya merubah opsi jawaban maka ada kemungkinan soal menjadi bias karena tidak fokus pada kemampuan mahasiswa dalam membandingkan. Oleh karena itu, revisi dilakukan dengan merubah $x = 1$ m menjadi $x = -1$ m.

Suatu benda ($m = 2$ kg) mula-mula dalam keadaan diam lalu bergerak disepanjang garis lurus dibawah pengaruh gaya yang berubah sebagai fungsi posisi mengikuti persamaan:

$$F(x) = 2x-4$$

 Seluruh besaran dalam satuan SI. Perbandingan kuadrat laju sesaat benda pada posisi $x = 0$, $x = -1$ m, dan $x = 4$ m adalah...

(A) $v_{x=0}^2 < v_{x=-1}^2 < v_{x=4}^2$
 (B) $v_{x=0}^2 = v_{x=-1}^2 < v_{x=4}^2$
 (C) $v_{x=0}^2 < v_{x=-1}^2 = v_{x=4}^2$
 (D) $v_{x=0}^2 = v_{x=4}^2 = 0 > v_{x=-1}^2$
 (E) $v_{x=0}^2 = v_{x=4}^2 = 0 < v_{x=-1}^2$

Gambar 1. Soal Nomor 8 Setelah Direvisi

Balok yang berada di atas permukaan licin mula-mula dalam keadaan diam di titik A kemudian ditarik ke kiri dengan gaya yang meningkat secara konstan hingga ke titik B. Dari titik B gaya dipertahankan konstan hingga titik C. Setelah di titik C, gaya dibuat sedemikian rupa sehingga balok mengalami perlambatan secara konstan hingga berhenti di titik D. Perhatikan ilustrasi berikut!

Disepakati bahwa tanda positif (+) merepresentasikan arah vektor ke kanan, dan sebaliknya. Jika jarak tiap segmen sama, manakah gaya sebagai fungsi posisi dalam representasi grafik yang paling sesuai dengan informasi tersebut?

Gambar 2. Soal Nomor 13 Setelah Direvisi

Soal nomor 13 pada **Gambar 2** merupakan soal yang menguji keterampilan kognitif merepresentasikan ulang (*re-representation*). Pada soal tersebut mahasiswa diminta untuk menentukan grafik $F(x)$ yang cocok sesuai kasus yang diberikan. Sebelum di revisi, seluruh opsi jawaban menunjukkan bahwa posisi (x) benda selalu bertambah. Meskipun berdasarkan perubahan magnitudo gaya memiliki jawaban yang benar, namun hal tersebut keliru karena tidak sesuai dengan konteks permasalahan, dimana benda selalu bergerak ke kiri. Oleh karena itu, soal tersebut direvisi (**Gambar 2**) dengan mengubah seluruh opsi jawaban. Seluruh opsi jawaban diubah dengan arah kurva pada sumbu x selalu ke berkurang/ke kiri.

3.2. Evaluasi Validitas dan Reliabilitas Instrumen Pada Uji Empiris

Analisis validitas butir soal dalam penelitian ini dilakukan dengan mengkaji korelasi biserial poin (r_{pbis}) antara masing-masing butir dan skor total tes. Hasilnya menunjukkan bahwa seluruh 15 butir soal memiliki nilai r_{pbis} yang signifikan ($p < 0,001$) dengan rentang nilai antara 0,355 hingga 0,666 (**Tabel 3**), sehingga seluruh butir dinyatakan valid. Temuan ini konsisten dengan penelitian sebelumnya yang menyatakan bahwa butir dengan $r_{pbis} \geq 0,30$ dapat dianggap memiliki daya diskriminasi yang baik untuk membedakan peserta sesuai dengan kemampuan yang diukur [29], [30].

Tabel 3. Analisis Validitas Butir Soal

Korelasi butir soal dengan skor total	r_{pbis}	p	Keterangan
Butir 1	0,605***	< ,001	Valid
Butir 2	0,534***	< ,001	Valid
Butir 3	0,666***	< ,001	Valid
Butir 4	0,566***	< ,001	Valid
Butir 5	0,355***	< ,001	Valid
Butir 6	0,464***	< ,001	Valid
Butir 7	0,421***	< ,001	Valid
Butir 8	0,486***	< ,001	Valid
Butir 9	0,541***	< ,001	Valid
Butir 10	0,522***	< ,001	Valid
Butir 11	0,455***	< ,001	Valid
Butir 12	0,449***	< ,001	Valid
Butir 13	0,424***	< ,001	Valid
Butir 14	0,523***	< ,001	Valid
Butir 15	0,455***	< ,001	Valid

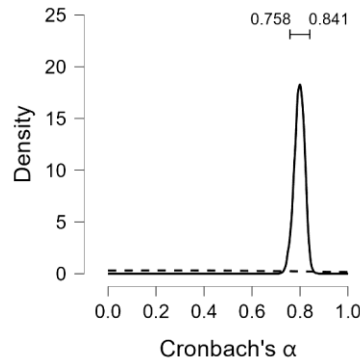
*** $p < 0,001$

Butir dengan korelasi tertinggi ditemukan pada butir 3 ($r_{pbis} = 0,666$), yang menunjukkan kemampuan diskriminatif sangat kuat, sedangkan butir 1 (0,605) dan butir 4 (0,566) juga menunjukkan hubungan yang tinggi dengan skor total. Nilai-nilai ini sejalan dengan laporan penelitian instrumen tes dalam konteks fisika oleh Ramadani, dkk. [31], Sandra, dkk. [32], dan Subair, dkk. [33], yang menunjukkan bahwa butir-butir dengan r_{pbis} signifikan merupakan kriteria penting agar instrumen mampu mengidentifikasi secara efektif peserta yang benar-benar memenuhi dan yang tidak memenuhi indikator pengukuran. Disisi lain, butir dengan nilai r_{pbis} terendah (butir 5, $r_{pbis} = 0,355$) tetap memenuhi kriteria validitas karena nilainya di atas *threshold* minimal yang direkomendasikan dalam analisis butir, yakni $r_{pbis} \geq 0,30$ [34]. Hal ini juga didukung oleh temuan penelitian lain yang menunjukkan bahwa nilai korelasi yang sedikit lebih rendah dapat muncul pada butir yang mengukur konsep yang kompleks atau memiliki variasi skenario kontekstual yang tinggi, tetapi masih dapat dipertahankan apabila secara teori memiliki kontribusi terhadap konstruk yang diukur [35].

Secara keseluruhan, sebagian besar butir (93,33%) memiliki nilai r_{pbis} di atas 0,40, yang menunjukkan bahwa instrumen ini secara konsisten mampu membedakan peserta berdasarkan tingkat pemahaman konsep mendalam pada topik Dinamika Partikel satu dimensi. Temuan ini mendukung gagasan dalam studi validitas instrumen pendidikan bahwa konsistensi nilai korelasi positif pada sebagian besar butir merupakan indikasi bahwa instrumen memiliki struktur valid internal yang kuat [36]. Secara statistik (Tabel 4), pada $n = 189$ ($df = 187$), korelasi sekitar $r = 0,143$ saja sudah signifikan pada taraf $\alpha = 0,05$ dua-arah, sedangkan $r_{pbis} = 0,30$ menghasilkan $t \approx 4,30$, sehingga jelas signifikan. Dengan demikian, karena setiap butir dari 15 soal memiliki r_{pbis} positif dan signifikan, apalagi mencapai $\geq 0,30$, maka tiap soal dapat dikatakan tidak hanya signifikan secara statistik, tetapi juga bermakna secara praktis dalam membedakan mahasiswa yang menguasai materi dan yang belum, sehingga kualitas perangkat tes menjadi lebih kuat untuk dipertahankan [37].

Tabel 4. Hasil Analisis Statistik Reliabilitas Skala Bayesian

		95% CI	
Koefesien	Posterior mean	Lower	Upper
Koefesien α	0.789	0.758	0.841



Gambar 3. Posterior Plots Analisis Reliabilitas Koefesien α

Selain validitas, reliabilitas instrumen dianalisis menggunakan pendekatan Bayesian. Hasil analisis menunjukkan koefisien Cronbach's α sebesar 0,789, dengan interval kepercayaan 95% antara 0,758–0,841. Distribusi posterior Cronbach's α pada **Gambar 3** memperlihatkan konsentrasi nilai di sekitar 0,789 dengan interval yang sempit, menandakan estimasi reliabilitas yang stabil dan tidak terlalu bervariasi. Nilai ini menunjukkan konsistensi internal instrumen yang baik, karena $\alpha \geq 0,7$ umumnya dianggap memadai untuk instrumen penelitian pendidikan [38], [39]. Hal ini menunjukkan bahwa butir-butir soal saling konsisten dan dapat diandalkan untuk mengukur pemahaman konsep mendalam mahasiswa. Dengan demikian, instrumen ini memenuhi syarat validitas dan reliabilitas, sehingga layak digunakan untuk menilai pemahaman konsep mendalam peserta pada topik dinamika partikel satu dimensi. Validitas yang tinggi dan reliabilitas yang baik menjadi dasar kuat untuk melanjutkan interpretasi hasil tes maupun pengembangan pembelajaran lebih lanjut.

3.3. Keterbatasan Penelitian dan Saran Penelitian Mendatang

Penelitian ini memiliki beberapa keterbatasan yang perlu dicermati, yaitu desainnya yang bersifat kuantitatif, non-eksperimental, deskriptif, dan *cross-sectional* sehingga temuan lebih menunjukkan kualitas awal instrumen daripada bukti yang lebih luas tentang kestabilan dan keefektifannya dalam berbagai konteks, meskipun telah diupayakan sumber data yang berasal dari berbagai universitas. Disamping itu, pengujian instrumen dalam studi ini berfokus pada validitas isi, validitas butir melalui r_{pbis} , dan reliabilitas konsistensi internal, sehingga aspek lain seperti validitas konstruk yang lebih mendalam, kestabilan skor dari waktu ke waktu, serta kinerja instrumen pada sampel yang lebih beragam belum sepenuhnya dieksplorasi. Oleh karena itu, penelitian mendatang disarankan untuk melibatkan sampel yang lebih luas serta menambahkan analisis psikometrik lanjutan agar diperoleh bukti validitas dan reliabilitas yang semakin komprehensif untuk mendukung penggunaan instrumen secara lebih luas. Instrumen penelitian ini juga dapat digunakan untuk mengidentifikasi pemahaman konsep mendalam untuk memetakan pengetahuan mahasiswa pada topik dinamika partikel 1 dimensi. Bagi peneliti yang memiliki fokus yang serupa, instrumen ini dapat digunakan untuk mengukur efektifitas implementasi desain pembelajaran terhadap pemahaman konsep mendalam melalui kegiatan eksperimental.

4. Kesimpulan

Berdasarkan hasil validasi ahli dan uji empiris, instrumen tes pemahaman konsep mendalam yang dikembangkan pada topik yang dikaji dapat dinyatakan memiliki kualitas yang baik dan layak digunakan sebagai alat ukur. Penilaian para ahli menunjukkan bahwa isi instrumen telah sesuai dengan konstruk yang hendak diukur, sementara masukan yang diberikan berhasil memperbaiki butir yang semula masih memiliki kelemahan konseptual dan representasional sehingga menjadi lebih tepat secara fisis maupun pedagogis. Hasil uji empiris juga memperlihatkan bahwa seluruh butir memiliki daya beda yang baik melalui nilai r_{pbis} yang signifikan, sehingga setiap soal mampu membedakan peserta didik berdasarkan tingkat penguasaan konsepnya.

Selain itu, analisis reliabilitas menunjukkan bahwa instrumen memiliki konsistensi internal yang baik, sehingga hasil pengukuran dapat dipercaya. Dengan demikian, instrumen ini dapat digunakan untuk menilai pemahaman konsep mendalam mahasiswa serta menjadi dasar yang kuat bagi evaluasi pembelajaran dan pengembangan instrumen lanjutan.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada seluruh ahli yang telah berkenan menjadi validator dan kepada mahasiswa yang telah berpartisipasi sebagai subjek dalam uji empiris instrumen soal. Ucapan terima kasih juga disampaikan kepada Lembaga Pengelola Dana Pendidikan (LPDP) atas dukungannya terhadap studi doctoral ini.

Data Pendukung

Data yang mendukung temuan penelitian ini, termasuk lembar validasi, kisi-kisi soal, data hasil validasi, dan data uji empiris telah disimpan secara terbuka di <https://zenodo.org/> dengan nomor akses <https://doi.org/10.5281/zenodo.19244415>. Peneliti lain dapat mengakses dan mengadaptasi instrumen penelitian secara terbuka.

Daftar Pustaka

- [1] M. R. A. Taqwa, P. Sinaga, E. Suhendi, D. Rochintaniawati, and H. F. Rahim, "Analysis of Outcome-Based Education Curriculum Implementation: Focus of Study on Classical Mechanics Course," *J. Phi J. Pendidik. Fis. dan Fis. Terap.*, vol. 11, no. 1, pp. 1–9, 2025, doi: 10.22373/p-jpft.v11i1.25857.
- [2] A. Puspitasari, F. Mufida, and B. Setiaji, "Analysis of the Influence of Analytical Mechanics Concept on the Development of Modern Physics," *J. Pendidik. dan Ilmu Fis.*, vol. 4, no. 1, pp. 60–88, 2024, doi: 10.52434/jpif.v4i1.4004.
- [3] K. Brown and H. Mathur, "Modified Newtonian Dynamics as an Alternative to the Planet Nine Hypothesis," *Astron. J.*, vol. 166, no. 4, p. 168, 2023, doi: 10.3847/1538-3881/acef1e.
- [4] J. Weber and T. Wilhelm, "Conceptual understanding of Newtonian dynamics in a comparative study of computational modeling and video motion analysis," in *Physics Education Research Conference Proceedings*, 2021, pp. 444–449. doi: 10.1119/perc.2021.pr.Weber.
- [5] A. Tongchai, M. D. Sharma, I. D. Johnston, K. Arayathanitkul, and C. Soankwan, "Consistency of students' conceptions of wave propagation: Findings from a conceptual survey in mechanical waves," *Phys. Rev. Spec. Top. Educ. Res.*, vol. 7, no. 020101, pp. 1–11, 2011, doi: 10.1103/PhysRevSTPER.7.020101.
- [6] S. Prihatiningtyas, M. H. Arrofi'uddin, and N. A. S. Pertiwi, "Learning media of physics-based on Google Sites with QR code on particle dynamics material," *J. Geliga Sains J. Pendidik. Fis.*, vol. 10, no. 2, pp. 134–143, 2022, doi: 10.31258/jgs.10.2.134-143.
- [7] Wahyudi, A. Sayhiral, and M. Zuhdi, "Menyingkap miskonsepsi gaya dan gerak: Studi kasus pada mahasiswa Pendidikan Fisika FKIP Universitas Mataram dan rekomendasi perbaikan," *J. Pendidikan, Sains, Geol. dan Geofis.*, vol. 6, no. 4, pp. 2433–2438, 2025, doi: 10.29303/goescienceed.v6i4.1458.
- [8] J. Weber and T. Wilhelm, "Contributing factors to the improvement of conceptual understanding in a computer-based intervention in Newtonian dynamics," *Phys. Rev. Phys. Educ. Res.*, vol. 20, no. 2, p. 20130, 2024, doi: 10.1103/PhysRevPhysEducRes.20.020130.
- [9] W. Y. Hwang, N. S. Chen, J. J. Dung, and Y. L. Yang, "Multiple representation skills and creativity effects on mathematical problem solving using a multimedia whiteboard system," *Educ. Technol. Soc.*, vol. 10, no. 2, pp. 191–212, 2007.
- [10] M. R. A. Taqwa, A. Suyudi, F. Sulman, and R. Faizah, "Kinematics students' conceptual understanding in mathematical and visual representations: Is it different?," *AIP Conf. Proc.*, vol. 2569, no. 1, p. 50005, Jan. 2023, doi: 10.1063/5.0112656.

- [11] T. A. Romansyah and M. R. A. Taqwa, "Are there differences in students' ability to solve vector problems in mathematical and visual representations?," *AIP Conf. Proc.*, vol. 2569, no. 1, p. 50022, Jan. 2023, doi: 10.1063/5.0112655.
- [12] D. Hestenes, M. Wells, and G. Swackhamer, "Force concept inventory," *Phys. Teach.*, vol. 30, no. 3, pp. 141–158, 1992, doi: 10.1119/1.2343497.
- [13] L. R. Dewi, W. Setyarsih, and L. Rohmawati, "Analisis Kualitas Instrumen Three-Tier Diagnostic Test Materi Dinamika Partikel," *J. Inov. Pendidik. Fis.*, vol. 5, no. 3, pp. 193–195, 2016.
- [14] M. R. D. Ernanda, S. Rizqiah, S. A. Binta, Y. Oktarisa, and T. G. S. Suryana, "Pengembangan Instrumen Tes Diagnostik Tingkat Konsepsi Pada Materi Kinematika," *J. Nebul.*, vol. 1, no. 2, pp. 56–60, 2024.
- [15] A. Fadllan, W. Y. Prawira, Arsini, and Hartono, "Analysis of students' misconceptions on mechanics using three-tier diagnostic test and clinical interview," *J. Phys. Conf. Ser.*, vol. 1170, no. 012027, 2019, doi: 10.1088/1742-6596/1170/1/012027.
- [16] E. Sulistri and Lisdawati, "Using three-tier test to identify the quantity of student that having misconception on newton's law of motion concept," *J. Ilmu Pendidik. Fis.*, vol. 2, no. 1, pp. 4–6, 2017.
- [17] L. W. Anderson and D. R. Krathwohl, *A taxonomy for learning, teaching and assessing: A revision of bloom's taxonomy of educational objectives: Complete edition*. New York: Longman, 2001.
- [18] L. Herayanti, H. Habibi, and B. A. Sukroyanti, "Development of inquiry-based teaching materials to improve physics teacher's conceptual understanding," *J. Penelit. Pendidik. IPA*, vol. 8, no. 6, pp. 3110–3116, 2022, doi: 10.29303/jppipa.v8i6.2543.
- [19] S. Ainsworth, "DeFT: A conceptual framework for considering learning with multiple representations," *Learn. Instr.*, vol. 16, pp. 183–198, 2006, doi: 10.1016/j.learninstruc.2006.03.001.
- [20] J. van der Meij and T. de Jong, "Supporting students' learning with multiple representations in a dynamic simulation-based learning environment," *Learn. Instr.*, vol. 16, pp. 199–212, 2006, doi: 10.1016/j.learninstruc.2006.03.007.
- [21] E. Almanasreh, R. Moles, and T. F. Chen, "Evaluation of methods used for estimating content validity," *Res. Soc. Adm. Pharm.*, vol. 15, no. 2, pp. 214–221, 2019, doi: 10.1016/j.sapharm.2018.03.066.
- [22] Jonathan P Bonnet, "Content and Face Validation of the Lifestyle Medicine Assessment," *Am. J. Lifestyle Med.*, vol. 18, no. 2, pp. 252–259, Mar. 2024, doi: 10.1177/15598276231214712.
- [23] D. F. Polit, C. T. Beck, and S. V Owen, "Is the CVI an acceptable indicator of content validity? Appraisal and recommendations," *Res. Nurs. Health*, vol. 30, no. 4, pp. 459–467, Aug. 2007, doi: <https://doi.org/10.1002/nur.20199>.
- [24] D. F. Polit and C. T. Beck, "The content validity index: Are you sure you know what's being reported? critique and recommendations," *Res. Nurs. Health*, vol. 29, no. 5, pp. 489–497, Oct. 2006, doi: 10.1002/nur.20147.
- [25] J. D. Brown, "Statistics corner: Questions and answers about language testing statistics: Point-biserial correlation coefficients," *Shiken JLT Test. Evolution SIG Newsl.*, vol. 5, no. 3, pp. 13–17, 2001.
- [26] D. Kornbrot, "Point biserial correlation," in *Wiley StatsRef: Statistics Reference Online* (eds N. Balakrishnan, T. Colton, B. Everitt, W. Piegorsch, F. Ruggeri and J.L. Teugels), 2014, pp. 1–3. doi: 10.1002/9781118445112.stat06227.
- [27] J. M. Pfadt, D. van den Bergh, K. Sijtsma, and E. J. Wagenmakers, "A tutorial on Bayesian single-test reliability analysis with JASP," *Behav. Res. Methods*, vol. 55, no. 3, pp. 1069–1078, 2023, doi: 10.3758/s13428-021-01778-0.
- [28] M. R. Lynn, "Determination and quantification of content validity," *Nursing Research*, vol. 35, no. 6. Lippincott Williams & Wilkins, US, pp. 382–385, 1986. doi: 10.1097/00006199-198611000-00017.
- [29] C. A. Mertler, *Introduction to educational research*. Sage publications, 2024.
- [30] D. Muijs, *Doing quantitative research in education with IBM SPSS statistics*. Sage Publications Ltd, 2022.
- [31] R. D. Ramadani, Y. Yennita, and E. Ernidawati, "Pengembangan instrumen tes berbasis literasi dains siswa SMP pada materi getaran dan gelombang," *Silampari J. Pendidik. Ilmu Fis.*, vol. 6, no. 1, pp. 25–34, 2024, doi: 10.31540/sjpif.v6i1.2428.

- [32] R. O. Sandra, M. Maison, and D. A. Kurniawan, "Pengembangan instrument miskonsepsi menggunakan dreamweaver berbasis web pada materi tekanan berformat five-tier," *J. Fis. Fis. Sains dan Apl.*, vol. 7, no. 1, pp. 22–28, 2022, doi: 10.35508/fisa.v7i1.6575.
- [33] M. Subair, S. S. Sari, and Jasruddin, "Pengembangan instrumen tes hasil belajar fisika kelas XI SMA Negeri Pangkajene," *J. Sains dan Pendidik. Fis.*, vol. 15, no. 1, pp. 32–35, 2019, doi: 10.35580/jspf.v15i1.9404.
- [34] M. J. Allen and W. M. Yen, *Introduction to measurement theory*. California: Brooks/Cole Publishing Company, 1979.
- [35] R. F. DeVellis and C. T. Thorpe, *Scale development: Theory and applications*. Sage publications, 2021.
- [36] J. W. Best and J. V Kahn, *Research in education: Tenth edition*. Harlow: Pearson Education, 2014.
- [37] D. Kaplan, B. Zumbo, and A. Rupp, "Responsible modeling of measurement data for appropriate inferences: Important advances in reliability and validity," in *The SAGE Handbook of Quantitative Methodology for the Social Sciences*, Sage. doi: 10.4135/9781412986311.n4.
- [38] K. S. Taber, "The use of Cronbach's Alpha when developing and reporting research instruments in science education," *Res. Sci. Educ.*, vol. 48, no. 6, pp. 1273–1296, 2018, doi: 10.1007/s11165-016-9602-2.
- [39] P. Govindasamy, T. M. Cumming, and N. Abdullah, "Validity and reliability of a needs analysis questionnaire for the development of a creativity module," *J. Res. Spec. Educ. Needs*, vol. 24, no. 3, pp. 637–652, Jul. 2024, doi: <https://doi.org/10.1111/1471-3802.12659>.

Apendiks A: Informasi Rancangan Instrumen

Nomor soal	Aspek berpikir	Indikator soal	Representasi (stimulus soal → opsi jawaban)
1	Menyontohkan (<i>exemplifying</i>)	Diberikan informasi kecepatan sebagai fungsi waktu dalam representasi grafik, mahasiswa mampu mencontohkan fenomena gerak akibat gaya konstan.	Grafik → verbal-simbolik
2	Menyontohkan (<i>exemplifying</i>)	Diberikan informasi kecepatan dan posisi sebagai fungsi waktu dalam representasi simbolik/matematis, mahasiswa mampu mencontohkan fenomena gerak akibat gaya konstan.	Verbal → matematis-simbolik.
3	Menyontohkan (<i>exemplifying</i>)	Diberikan informasi diagram gerak, mahasiswa mampu mencontohkan fenomena gerak akibat gaya konstan.	Diagram gerak (<i>pictorial</i>) → verbal-simbolik
4	Mengklasifikasi (<i>Classifying</i>)	Diberikan informasi kecepatan sebagai fungsi waktu dari beberapa benda dalam representasi grafik, mahasiswa mampu mengklasifikasikan keadaan gaya yang bekerja (nol, konstan tidak nol, dan tidak konstan).	Grafik → verbal-simbolik
5	Mengklasifikasi (<i>Classifying</i>)	Diberikan informasi kecepatan dan percepatan benda selama bergerak dalam representasi verbal, mahasiswa mampu mengklasifikasikan keadaan gaya yang bekerja (nol, konstan tidak nol, dan tidak konstan).	Verbal → verbal-simbolik
6	Mengklasifikasi (<i>Classifying</i>)	Diberikan informasi percepatan sebagai fungsi waktu dari beberapa benda dalam representasi grafik, mahasiswa mampu mengklasifikasikan keadaan gaya yang bekerja (nol, konstan tidak nol, dan tidak konstan).	Grafik → verbal-simbolik
7	Membandingkan (<i>Comparing</i>)	Diberikan informasi gaya sebagai fungsi waktu dari beberapa benda dalam representasi grafik, mahasiswa mampu membandingkan besar perubahan kecepatan.	Grafik → simbolik
8	Membandingkan (<i>Comparing</i>)	Diberikan informasi gaya sebagai fungsi posisi dalam representasi simbolik (matematis), mahasiswa mampu membandingkan besar kuadrat kecepatan pada posisi-posisi tertentu.	Matematis-simbolik → simbolik
9	Membandingkan (<i>Comparing</i>)	Diberikan informasi posisi benda tiap detik dalam representasi visual (diagram gerak), mahasiswa mampu membandingkan besar resultan gaya pada posisi-posisi tertentu.	Diagram gerak (<i>pictorial</i>) → simbolik
10	Menjelaskan (<i>Explaining</i>)	Diberikan informasi gaya yang bekerja pada benda tiap detik dalam representasi grafik, mahasiswa mampu menjelaskan perubahan besar dan arah gaya.	Grafik → verbal-simbolik
11	Menjelaskan (<i>Explaining</i>)	Diberikan informasi kecepatan benda tiap detik dalam representasi tabel, mahasiswa mampu menjelaskan perubahan besar dan arah gaya.	Tabel → verbal
12	Menjelaskan (<i>Explaining</i>)	Diberikan informasi gaya yang bekerja pada benda dan kecepatan tiap detik dalam representasi grafik, mahasiswa mampu menjelaskan hubungan antara kecepatan dan gaya.	Grafik → verbal
13	Merepresentasikan Ulang (<i>Rerepresenting</i>)	Diberikan informasi gaya yang bekerja pada benda yang berubah sebagai fungsi posisi dalam representasi visual (gambar) dan verbal, mahasiswa dapat merepresentasikan ulang informasi $F(x)$ dalam representasi grafik.	Verbal-visual → grafik
14	Merepresentasikan Ulang (<i>Rerepresenting</i>)	Diberikan informasi gaya yang bekerja pada benda yang berubah sebagai fungsi waktu dalam representasi grafik, mahasiswa dapat merepresentasikan perubahan besar percepatan secara verbal.	Grafik → verbal-simbolik
15	Merepresentasikan Ulang (<i>Rerepresenting</i>)	Diberikan informasi gaya dan energi potensial sebagai fungsi posisi dalam representasi matematis, mahasiswa dapat merepresentasikan gaya dan energi potensial dalam representasi grafik.	Matematis-simbolik → grafik

