

---

## **Pengembangan instrumen penilaian untuk mengukur kemampuan berpikir kritis pada materi dinamika partikel secara *online***

### **Safira**

Program Studi Pendidikan Fisika, Universitas Lampung, Indonesia  
Surat-e: sapsafira491@gmail.com

### **Eko Suyanto**

Program Studi Pendidikan Fisika, Universitas Lampung, Indonesia  
Surat-e: eko.suyanto@fkip.unila.ac.id

### **I Dewa Putu Nyeneng**

Program Studi Pendidikan Fisika, Universitas Lampung, Indonesia  
Surat-e: idewaputunyenengms@gmail.com

**Abstrak.** Peserta didik SMA dibiasakan berpikir sederhana, bahkan pada pelajaran sains, seringkali konten materi tidak tersampaikan. Instrumen penilaian yang berfokus pada penerapan Hukum Newton I, II, III, khususnya instrumen yang tidak menggunakan perhitungan angka, masih sulit ditemukan. Penelitian ini bertujuan mengembangkan instrumen penilaian jenis uraian untuk mengukur kemampuan berpikir kritis materi dinamika partikel, serta menguji kebakuan instrumen penilaian. Metode penelitian menggunakan *Research and Development*, dengan langkah: mendefinisikan konstruk dan tujuan, menentukan format, konstruksi butir soal, pembuatan panduan penilaian, validasi teoritik, uji lapangan, analisis kebakuan, dan revisi. Uji coba pada 67 peserta didik yang telah mempelajari dinamika partikel. Hasil penelitian berjumlah 8 butir soal valid, reliabilitas 0,86 kriteria sangat tinggi, taraf kesukaran: 1 sukar dan 9 sedang, daya pembeda: 2 baik sekali, 2 baik, 4 cukup, dan 2 tanpa daya pembeda. Instrumen penilaian ini dapat mengukur, sekaligus melatih peserta didik dalam meningkatkan kemampuan berpikir kritis.

**Kata kunci:** Asesmen Fisika, SMA, Tes Essay

**Abstract.** High school students are accustomed to simple thinking, even in science lessons, often the content of the material is not conveyed. Assessment instruments that focus on applying Newton's Laws, especially without numerical calculations, are still hard to find. The research aims to develop an essay assessment instrument to measure critical thinking skills in particle dynamics and to test the standard of the items. The research method uses Research and Development, steps: defining the construct, determining the format, making an assessment guide, theoretical validation, field testing, and standard analysis. Trial of 67 students who have studied particle dynamics. The results obtained 8 valid items, with reliability 0.86 very high, level of difficulty: 1 difficult and 9 medium, discrimination power: 2 very good, 2 good, 4 fair, and 2 without discrimination. This instrument can measure, as well as train students in improving critical thinking skills.

**Keywords:** Physics Assessment, High-School, Essay Test

---

## **I. Pendahuluan**

Perubahan peradaban dunia yang berkembang cepat, akan berdampak dalam berbagai bidang seperti teknologi dan dunia kerja, sehingga semakin kompleksnya dunia kerja diperlukan tuntutan kemampuan masa depan yang disesuaikan dengan abad 21. Kemampuan pada abad 21 yang harus dimiliki adalah kemampuan berpikir kritis [1], [2], kemampuan tersebut melatih dalam bekerja secara efektif dan kompetitif di pasar global. Selain itu, berpikir kritis juga digunakan untuk pengambilan keputusan dalam penyelesaian masalah

secara rasional dan ilmiah [3], [4]. Tanpa kemampuan berpikir kritis menyebabkan penanggapan informasi yang keliru, sehingga berdampak negatif dalam pengambilan keputusan [5].

Pengembangan dan transfer kemampuan berpikir kritis diakui merupakan tujuan terpenting khususnya dunia pendidikan. Kemampuan berpikir kritis ialah tujuan utama pendidikan nasional, khususnya pada pendidikan menengah maupun tinggi [6], [7], [8].

Faktanya peserta didik pada SMA dibiasakan berpikir sederhana, disebabkan minimnya kompetensi guru dalam menerapkan kemampuan berpikir kritis [9], [10], [11]. Sedikit sekali yang diketahui oleh guru tentang cara terbaik untuk melatih peserta didik dalam berpikir kritis [12]. Bahkan pada mata pelajaran yang sangat diperlukan untuk berpikir kritis seperti sains, seringkali konten materi tidak tersampaikan [13]. Kenyataan ini menggambarkan bahwa peserta didik belum sepenuhnya menerapkan kemampuan berpikir kritis secara efektif dan berkelanjutan.

Berdasarkan hasil studi peringkat PISA tahun 2018, Indonesia mengalami penurunan dibandingkan tahun 2015, yaitu 9 terbawah dari 80 negara pada penilaian kinerja sains, nilai rata-rata kinerja sains peserta didik Indonesia sebesar 396 (dari 1000) [14]. Sebanyak 41,4% peserta didik berada di level 1a, dimana peserta didik hanya mampu mengidentifikasi penjelasan fenomena ilmiah secara sederhana dengan tingkat kognitif yang rendah. Persentase peserta didik yang berprestasi dibidang sains hanya sebesar 1,7%, yang berarti mereka mahir di Level 4 dan 5, sedangkan level 6 sebesar 0%.

Instrumen penilaian PISA terdiri dari pertanyaan-pertanyaan mengenai fenomena kehidupan nyata yang bertujuan untuk membangun argumen peserta didik dan membutuhkan penalaran berpikir kritis yang mendalam. Data PISA Indonesia tahun 2018 membuktikan bahwa kemampuan kinerja sains peserta didik dalam mengembangkan kemampuan berpikir kritis masih jauh dibawah rata-rata. Hal ini disebabkan kurangnya pelatihan dalam memecahkan soal-soal tingkat tinggi, khususnya dalam berpikir kritis.

Fakta lapangan analisis pendahuluan, guru sering menggunakan instrumen penilaian kategori C1-C3 dalam kategori *LOTS (Low Order Thinking Skill)*, dan hanya menggunakan level C4 dalam kategori *HOTS (Higher Order Thinking Skill)*. Walaupun peserta didik sudah cukup dilatih, akan tetapi guru masih menemui kesulitan dalam pengaplikasian level kognitif C5 dan C6, level tersebut termasuk kedalam indikator dalam berpikir kritis.

Selain itu, guru jarang menggunakan instrumen penilaian dalam bentuk tulisan dengan tujuan membiasakan peserta didik untuk dapat berpikir kritis [11], [16], [17]. Peserta didik akan mendapatkan pengalaman serta pengetahuan baru dari hasil menganalisis dan menyelidiki instrumen penilaian yang diberikan oleh guru [18]. Instrumen penilaian berpikir kritis saat ini sebagian besar bersifat umum, sehingga kurangnya instrumen yang bersifat spesifik. Instrumen penilaian kemampuan berpikir kritis, secara spesifik pada mata pelajaran fisika belum banyak yang mengembangkan [7]. Berkaitan dengan pernyataan tersebut, penting untuk mengembangkan suatu instrumen yang dapat mengukur kemampuan berpikir kritis, khususnya bidang fisika di SMA.

Peserta didik umumnya, sulit memahami materi mengenai dinamika partikel, terutama dalam memecahkan suatu permasalahan pada konsep tersebut [19]. Peserta didik menganggap bahwa teori-teori fisika yang mereka pelajari hanyalah sekadar rumus yang bisa dimanipulasi, tanpa memikirkan mengenai potensi yang dapat diterapkan dalam kehidupan. Hal tersebut sesuai dengan penelitian [20] mengungkapkan bahwa peserta didik mengalami kebingungan dalam memahami konsep dinamika partikel yang disebabkan peserta didik kurang membiasakan dalam menganalisis masalah yang berkaitan dalam penguasaan konsep.

Hal ini sejalan dengan hasil studi lapangan menggunakan angket *google form* yang disebarkan kepada peserta didik kelas XI IPA disalah satu sekolah swasta Bandar Lampung, yaitu sebanyak 65,7% peserta didik menjawab keliru perihal materi dinamika partikel yang telah mereka pelajari. Peserta didik menganggap bahwa konsep gerak rotasi, hukum hooke, titik berat, elastisitas, torsi, momen inersia, dan momen gaya, serta kesetimbangan benda tegar merupakan bagian dari dinamika partikel. Peserta didik mengalami permasalahan dalam menjawab konsep materi dinamika partikel, sejalan dengan hasil studi pendahuluan yang menjabarkan mengenai materi dinamika partikel yang dirasakan peserta didik, antara lain: sebanyak 61,4% tidak menyukai, 70% merasa kesulitan, dan 75,7% kurang memahami konsep secara keseluruhan. Berdasarkan hasil studi pendahuluan dapat disimpulkan bahwa tidak ada ketertarikan terhadap materi ini, menyebabkan peserta didik merasa kesulitan memahami materi, sehingga menimbulkan ketidakpahaman terhadap konsep yang diajarkan.

Berdasarkan penjabaran di atas, maka diperlukan instrumen penilaian berpikir kritis materi dinamika partikel yang teruji kebakuan tiap butir soal. Oleh karena itu, dilakukanlah penelitian pengembangan yang

berjudul “Pengembangan Instrumen Penilaian Untuk Mengukur Kemampuan Berpikir Kritis Pada Materi Dinamika Partikel Secara *Online*”.

## II. Metode Penelitian

*Research and Development (R&D)* dipilih sebagai metode penelitian ini, yang mengacu pada langkah penelitian pengembangan hasil asimilasi [7], [21], [22] yang terdiri dari: mendefinisikan konstruk dan merumuskan tujuan, menentukan format instrumen penilaian berpikir kritis, konstruksi butir soal berpikir kritis, pembuatan panduan penilaian, validasi teoritik dan uji lapangan, uji validasi empirik, tingkat reliabilitas, daya pembeda dan taraf kesukaran, terakhir revisi butir soal. Pengembangan ini bertujuan untuk menghasilkan instrumen penilaian jenis uraian yang dapat mengukur kemampuan berpikir kritis dan menguji kebakuan butir soal.

Kemampuan berpikir kritis yang ditargetkan dalam instrumen penilaian jenis uraian dipilih untuk digunakan pada materi dinamika partikel. Instrumen penilaian ini berfokus pada 3 aspek kemampuan berpikir kritis, meliputi: pengujian hipotesis, analisis kemungkinan dan ketidakpastian, serta pengambilan keputusan dan pemecahan masalah. Peneliti menggunakan 6 Indikator kemampuan berpikir kritis, yang dijabarkan menurut [7], pada Tabel 1.

**Tabel 1.** Aspek dan Indikator Berpikir Kritis Yang Dikembangkan

No.	Aspek Berpikir Kritis	Indikator Berpikir Kritis – Spesifik Dalam Bidang Sains
1.	Pengujian Hipotesis	<ul style="list-style-type: none"> <li>• menginterpretasikan hubungan antar variabel</li> <li>• mengakui kebutuhan akan lebih banyak informasi dalam menarik kesimpulan</li> </ul>
2.	Analisis kemungkinan dan ketidakpastian	<ul style="list-style-type: none"> <li>• memprediksi kemungkinan suatu peristiwa</li> <li>• memahami kebutuhan data tambahan dalam menentukan keputusan</li> </ul>
3.	Pemecahan masalah dan pengambilan keputusan	<ul style="list-style-type: none"> <li>• mengenali pilihan terbaik di antara beberapa pilihan dalam menangani masalah</li> <li>• mengevaluasi solusi untuk masalah dan membuat keputusan yang kuat</li> </ul>

Angket analisis kebutuhan peserta didik (*google form*), angket wawancara, angket uji ahli, dan instrumen penilaian berpikir kritis dinamika partikel merupakan instrumen penilaian yang digunakan pada penelitian ini. Pengisian angket penelitian pendahuluan dilaksanakan oleh peserta didik dan guru. Pengisian angket uji kelayakan dilakukan oleh ahli atau *expert*, yaitu 2 dosen FKIP Universitas Lampung yang ahli dalam bidang penelitian ini dan 1 guru fisika sebagai praktisi.

## III. Hasil Penelitian dan Pembahasan

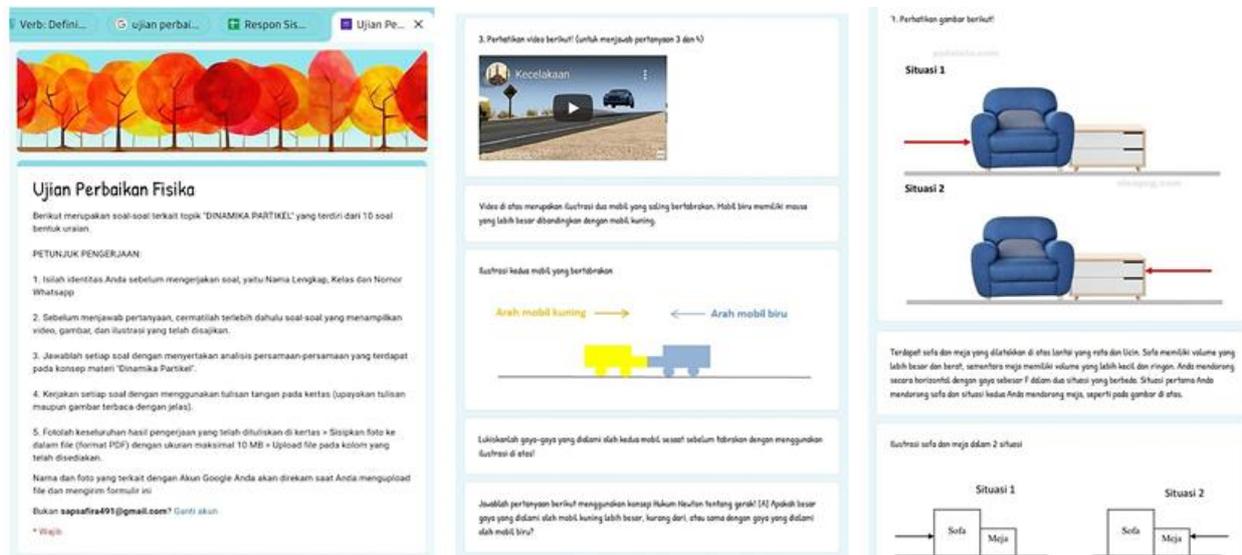
Instrumen penilaian, sebelum diujicobakan hendaknya diuji kevalidan terlebih dahulu kepada ahli substansi, konstruk, dan bahasa agar data yang diperoleh dapat dipertanggungjawabkan, seperti pada Tabel 2.

**Tabel 2.** Hasil Validasi Ahli Secara Keseluruhan

No	Aspek	Validator			Nilai Rata-rata	Kriteria
		1	2	3		
1.	Substansi	79%	88%	83%	83%	SV
2.	Konstruksi	95%	90%	95%	93%	SV
3.	Bahasa	100%	88%	94%	94%	SV
Nilai Rata-Rata					90%	SV

Kriteria : SV (Sangat Valid)

Hasil validasi ahli secara keseluruhan diperoleh nilai hasil validasi ahli dari ketiga aspek yaitu sebesar 90% yang dikategorikan pada kriteria sangat valid. Instrumen penilaian yang telah divalidasi oleh ahli kemudian diujikan kepada 67 peserta didik kelas X MIPA SMA Al Kautsar yang telah mempelajari materi dinamika partikel. Uji terbatas ini dimulai pada tanggal 25 April 2021. Berikut contoh produk instrumen penilaian berpikir kritis materi dinamika partikel secara *online* menggunakan *google form*, pada Gambar 1.



**Gambar 1.** Instrumen Penilaian Berpikir Kritis Materi Dinamika Partikel Secara *Online*

Berdasarkan hasil penggunaan instrumen penilaian kemampuan berpikir kritis materi dinamika partikel, yang ditinjau dari 3 aspek keseluruhan, yang digunakan pada 67 peserta didik X MIPA SMA Al Kautsar, memiliki nilai rata-rata sebesar setengah dari nilai maksimal. Nilai rata-rata 10 butir soal yaitu 53 (dari 100). Informasi terkait pengerjaan instrumen penilaian dapat dilihat dari hasil uji lapangan pada Tabel 3.

**Tabel 3.** Analisis Hasil Uji Coba Lapangan

No	Aspek Berpikir Kritis	Nomor Soal	Persentase	Rata-Rata Persentase Per-Aspek
1.	Pengujian Hipotesis	5	52%	53%
		6	60%	
		3	49%	
		4	52%	
2.	Analisis Kemungkinan dan Ketidakpastian	1	44%	51%
		2	68%	
		8	61%	
		10	30%	
3.	Pemecahan Masalah dan Pengambilan Keputusan	7	46%	54%
		9	61%	

Pada Tabel 4, aspek pemecahan masalah dan pengambilan keputusan memiliki persentase terbesar yaitu 54%, selanjutnya aspek pengujian hipotesis sebesar 53%, dan aspek analisis kemungkinan dan ketidakpastian memiliki persentase terkecil yaitu sebesar 51%. Jika dilihat pada hasil persentase 3 aspek berpikir kritis maka perbandingannya terlihat sama. Namun, dari ketiga aspek yang memiliki persentase tertinggi adalah aspek pemecahan masalah dan pengambilan keputusan. Aspek ini dibutuhkan untuk mengidentifikasi dan menghasilkan tindakan kreatif dan alternatif untuk membuat keputusan dan memecahkan masalah. Indikator pada aspek ini, terdiri dari: mengidentifikasi keputusan terbaik di antara

beberapa alternatif dalam memecahkan masalah yang tertera pada nomor 9, dan mengevaluasi solusi masalah dan membuat keputusan yang kuat pada soal nomor 7.

Persentase kebenaran dalam menjawab soal nomor 7 lebih besar yaitu 61% dibandingkan dengan soal nomor 9 yaitu 46%. Pada soal nomor 7, 61% peserta didik dapat menganalisis dan menjawab pertanyaan dengan menyajikan penyelesaian secara matematis dalam bentuk persamaan, sehingga peserta didik dapat mengambil keputusan dalam memecahkan masalah pada soal nomor 7. Namun, sebagian besar peserta didik kesulitan dalam melukiskan gaya-gaya pada ilustrasi yang telah disajikan dengan presisi. Pada soal nomor 9, 46% peserta didik dapat menganalisis dan menjawab pertanyaan secara logika dengan melukiskan gaya pada bola. Namun, sebagian besar peserta didik kesulitan untuk membuktikan bahwa percepatan tersebut adalah percepatan gravitasi yang berlaku disetiap bola pada soal nomor 9.

Aspek analisis kemungkinan dan ketidakpastian memiliki persentase terkecil yaitu sebesar 51%. Aspek analisis kemungkinan dan ketidakpastian dibutuhkan untuk dapat memprediksi kemungkinan keberhasilan dan kegagalan dalam pengambilan keputusan. Terdapat 2 indikator berpikir kritis yang digunakan, yaitu indikator memprediksi kemungkinan suatu peristiwa yang memiliki persentase rata-rata sebesar 52% yang tertera pada nomor 1 dan 8, dan mengidentifikasi asumsi yang memiliki persentase rata-rata sebesar 49% yang tertera pada soal nomor 2 dan 10. Persentase indikator berpikir kritis memprediksi kemungkinan suatu peristiwa lebih besar dibandingkan mengidentifikasi asumsi.

Persentase kebenaran dalam menjawab soal pada nomor 1 dan 10 lebih besar dibandingkan pada soal nomor 2 dan 8. Berdasarkan hasil analisis validitas menggunakan ANATES, soal 2 dan 8 dinyatakan tidak signifikan atau tidak valid disebabkan sebagian besar peserta didik pada kelompok bawah (asor) dapat menjawabnya dengan benar, sementara sebagian besar peserta didik pada kelompok unggulan kurang tepat dalam menjawab permasalahan. Sehingga soal nomor 2 dan 8 termasuk kategori yang tidak memiliki daya pembeda.

Soal nomor 10 memiliki persentase sangat kecil yaitu sebesar 30%, dibandingkan soal pada nomor 1 yaitu sebesar 44%. Sebanyak 30% peserta didik yang menjawab soal nomor 10, dapat menganalisis dan melukiskan gaya-gaya yang bekerja pada dua situasi tersebut. Namun, sebagian besar peserta didik kurang presisi dalam melukiskan gaya-gaya, serta kurang tepat dalam menganalisis secara matematis yaitu pembuktian persamaan, sehingga menyebabkan kekeliruan dalam mengidentifikasi asumsi dari aspek analisis kemungkinan dan ketidakpastian berpikir kritis.

Pada matriks rekapitulasi analisis butir soal, terdapat 7 kolom yang terdiri dari: No. (Nomor Butir Soal), TK (Tarf Kesukaran), DP (Daya Pembeda), Korelasi atau nilai validitas, kolom Signifikan, kolom V/TV (Valid/Tidak Valid), dan kolom Ket. (Keterangan) berupa SD (Soal Dipakai), STD (Soal Tidak Dipakai), serta SDD (Soal Dipakai Diperbaiki) seperti pada Tabel 4.

**Tabel 4.** Rekapitulasi Analisis Butir Soal

No.	DP	TK	Korelasi	Signifikan	V/TV	Ket
1.	0.283	Sedang	0.600	Signifikan	V	SDD
2.	-0.138	Sedang	0.066	Tidak Signifikan	TV	STD
3.	0.727	Sedang	0.641	Signifikan	V	SD
4.	0.844	Sedang	0.760	Sangat Signifikan	V	SD
5.	0.300	Sedang	0.577	Signifikan	V	SDD
6.	0.314	Sedang	0.665	Signifikan	V	SDD
7.	0.671	Sedang	0.816	Sangat Signifikan	V	SD
8.	-0.055	Sedang	0.156	Tidak Signifikan	TV	STD
9.	0.333	Sedang	0.631	Signifikan	V	SDD
10.	0.523	Sukar	0.644	Signifikan	V	SDD

Butir soal yang dapat langsung dipakai atau diterima adalah butir soal yang memiliki validitas  $\geq 0,576$  dengan taraf signifikan sebesar 5% berdasarkan analisis ANATES 4.0.5., taraf kesukaran direntang 0,31 – 0,70 dengan kategori sedang [23], dan daya pembeda direntang 0,40 – 1,00 dengan kategori baik dan baik sekali [24]. Hasil dari rekapitulasi analisis butir soal yang dapat langsung dipakai yaitu nomor 3, 4, dan 7.

Butir soal yang bisa dipakai namun perlu perbaikan adalah butir soal yang memiliki validitas  $\geq 0,576$  berdasarkan analisis ANATES 4.0.5., namun memiliki taraf kesukaran yang berada pada rentang 0,00 – 0,30 atau 0,71 – 1,00 dengan kategori mudah dan sukar [23], serta daya pembeda direntang 0,00 – 0,39 dengan kategori lemah dan cukup [24]. Hasil dari rekapitulasi analisis butir soal yang dapat dipakai namun perlu perbaikan berjumlah 5 butir soal, tepatnya pada nomor 1, 5, 6, 9, dan 10.

Butir soal yang harus dibuang adalah butir soal yang memiliki validitas  $< 0,576$  berdasarkan analisis ANATES 4.0.5., yang seharusnya disertai juga dengan nilai daya pembeda  $< 0$  atau bertanda negatif, kategori tidak ada daya pembeda. Tanda negatif pada nilai daya pembeda menunjukkan bahwa butir soal mampu dijawab oleh peserta didik berkemampuan rendah, namun yang berkemampuan tinggi tidak mampu menjawabnya. Seharusnya, soal mampu dijawab tepat oleh peserta didik yang berkemampuan tinggi dibandingkan yang berkemampuan rendah. Sehingga dapat diputuskan, soal yang harus dibuang adalah soal pada nomor 2 dan 8.

Berdasarkan hasil rekapitulasi analisis butir soal maka dapat diketahui hubungan validitas, taraf kesukaran dan daya pembeda. Hubungan nilai validitas dan daya pembeda butir soal searah, sesuai dengan hasil penelitian [22]. Sementara itu, nilai validitas dengan taraf kesukaran tidak menunjukkan hubungan yang searah, sesuai dengan hasil penelitian [25]. Nilai taraf kesukaran dengan daya pembeda tidak menunjukkan hubungan yang searah, sesuai dengan hasil penelitian [26].

Berdasarkan hasil penelitian sebelumnya, melalui instrumen penilaian, peserta didik dapat mengasah kemampuan dalam berpikir secara kritis. [27] dalam penelitiannya, menunjukkan bahwa dengan penggunaan instrumen penilaian HOTS uraian dapat meningkatkan kemampuan berpikir peserta didik. Penelitian [28] mengenai pengembangan instrumen penilaian HOTS fisika juga memperoleh hasil yang sama bahwa dengan menggunakan penilaian HOTS seperti tes *essay* dapat meningkatkan kemampuan berpikir peserta didik.

Sebagian besar instrumen penilaian berpikir kritis dibuat dalam format pilihan ganda, misalnya *California Critical Thinking Skills Test*, *Halpern Critical Thinking Assessment*, dan *ACT Collegiate Assessment*. Hal ini bertujuan untuk mengurangi bias jawaban peserta agar terkontrol dalam penilaiannya [29], [30]. Namun, orisinalitas pemikiran kritis peserta didik tidak akan keluar, karena jawaban telah dikendalikan [29], [31]. Instrumen penilaian berpikir kritis materi dinamika partikel dalam format uraian mengungkapkan orisinalitas pemikiran kritis peserta dengan mempertimbangkan bias jawaban melalui desain domain-spesifik dan indikator yang komprehensif. Instrumen penilaian ini dapat mengukur, sekaligus melatih peserta didik dalam meningkatkan kemampuan berpikir kritis.

#### IV. Kesimpulan

Kesimpulan penelitian ini, diperoleh 8 butir soal valid dengan kategori: 2 soal sangat signifikan (sangat valid) dan 6 soal signifikan (valid). Nilai reliabilitas instrumen penilaian sebesar 0,86 dengan kriteria sangat tinggi. Taraf kesukaran berkisar (0,298) – (0,680) dengan kategori: 9 soal sedang dan 1 sukar. Daya pembeda berkisar (-0,138) – (0,844) dengan kategori: 2 soal tanpa daya pembeda, 4 cukup, 2 baik, dan 2 baik sekali. Saran bagi peneliti lain, diharapkan dapat mengembangkan instrumen penilaian jenis uraian pada materi fisika lainnya, yang menekankan pada penerapan di kehidupan sehari-hari.

#### Ucapan Terimakasih

Ucapan terima kasih kepada 3 validator pada penelitian pengembangan ini, yaitu, Dr. Abdurrahman, M.Si., Dr. Doni Andra, M.Sc., dan Supardi, M.Pd. Terimakasih kepada Kepala Sekolah yang telah mengizinkan peneliti untuk menguji kebakuan instrumen penilaian, dan juga guru SMA Al Kautsar yang telah membantu dalam pelaksanaan penelitian ini, Nurazmi, S.Pd., dan Miftahul Hasanah, S.Pd.

#### Kepustakaan

- [1] M. Kang, H. Heo, L. Jo, J. Shin, and J. Seo, "Developing an educational performance indicator for new millennium learners," *J. Res. Technol. Educ.*, vol. 43, no. 2, pp. 157–170, 2010, doi: 10.1080/15391523.2010.10782567.

- [2] S. C. Kong, "Developing information literacy and critical thinking skills through domain knowledge learning in digital classrooms: An experience of practicing flipped classroom strategy," *Comput. Educ.*, vol. 78, pp. 160–173, 2014, doi: 10.1016/j.compedu.2014.05.009.
- [3] O. Sarigoz, "Assessment of the High School Students' Critical Thinking Skills," *Procedia - Soc. Behav. Sci.*, vol. 46, pp. 5315–5319, 2012, doi: 10.1016/j.sbspro.2012.06.430.
- [4] D. F. Halpern, *Thought and knowledge: An introduction to critical thinking, Fifth Edition*, New York : Psychology Press, 2014.
- [5] R. Stobaugh, *Assessing critical thinking in middle and high schools: Meeting the Common Core*, New York : Routledge, 2013.
- [6] F. Mabruroh and A. Suhandi, "Construction Of Critical Thinking Skills Test Instrument Related The Concept On Sound Wave," *J. Phys. Conf. Ser.*, vol. 812, p. 12056, 2017, doi: 10.1088/1742-6596/812/1/012056.
- [7] D. T. Tiruneh, M. De Cock, A. G. Weldeslassie, J. Elen, and R. Janssen, "Measuring Critical Thinking in Physics: Development and Validation of a Critical Thinking Test in Electricity and Magnetism," *Int. J. Sci. Math. Educ.*, vol. 15, no. 4, pp. 663–682, 2017, doi: 10.1007/s10763-016-9723-0.
- [8] W.-J. Shim and K. Walczak, "The Impact of Faculty Teaching Practices on the Development of Students' Critical Thinking Skills," *Int. J. Teach. Learn. High. Educ.*, vol. 24, no. 1, pp. 16–30, 2012.
- [9] T. Carlgren, "Communication, Critical Thinking, Problem Solving: A Suggested Course for All High School Students in the 21st Century," *Interchange*, vol. 44, no. 1–2, pp. 63–81, 2013, doi: 10.1007/s10780-013-9197-8.
- [10] S. C. Choy, "Reflective Thinking and Teaching Practices: a Precursor for Incorporating Critical Thinking Into the Classroom?," *Online Submiss.*, vol. 5, no. 1, pp. 167–182, 2012.
- [11] J. Sermeus, M. De Cock, and J. Elen, "Critical Thinking in Electricity and Magnetism : an Assessment Tool for Secondary School Students, " in *European Science Education Research Association (ESERA), Date: 2017/08/21-2017/08/25, Location: Dublin*, 2010, pp. 2–4.
- [12] P. K. Murphy, M. L. Rowe, G. Ramani, and R. Silverman, "Promoting Critical-Analytic Thinking in Children and Adolescents at Home and in School," *Educ. Psychol. Rev.*, vol. 26, no. 4, pp. 561–578, 2014, doi: 10.1007/s10648-014-9281-3.
- [13] L. M. Marin and D. F. Halpern, "Pedagogy for developing critical thinking in adolescents: Explicit instruction produces greatest gains," *Think. Ski. Creat.*, vol. 6, no. 1, pp. 1–13, 2011, doi: 10.1016/j.tsc.2010.08.002.
- [14] OECD, *PISA 2018 Results (Volume I): What Students Know and Can Do*, Paris : OECD Publishing, 2019.
- [15] OECD, *PISA 2018: Insights and Interpretations*, Paris : OECD Publishing, 2019.
- [16] W. Wartono, M. N. Hudha, and J. R. Batlolona, "How are the physics critical thinking skills of the students taught by using inquiry-discovery through empirical and theoretical overview?," *Eurasia J. Math. Sci. Technol. Educ.*, vol. 14, no. 2, pp. 691–697, Dec. 2018, doi: 10.12973/ejmste/80632.
- [17] T. Sugiarti, I. Kaniawati, and L. Aviyanti, "Development of Assessment Instrument of Critical Thinking in Physics at Senior High School," *J. Phys. Conf. Ser.*, vol. 812, p. 12018, 2017, doi: 10.1088/1742-6596/812/1/012018.
- [18] R. Herpiana, U. Rosidin, and A. Abdurrahman, "Development of Instruments to Train Critical and Creative Thinking Skills in Physics Assessment for High School Students' Learning," *J. Phys. Conf. Ser.*, vol. 1155, no. 1, 2019, doi: 10.1088/1742-6596/1155/1/012046.
- [19] J. Machado and M. Braga, "Secondary students' modelling conceptualisation in situations related to particle dynamics: a clinical perspective," *Int. J. Sci. Educ.*, vol. 40, no. 13, pp. 1606–1628, 2018, doi: 10.1080/09500693.2018.1494394.
- [20] I. A. D. Astuti, I. Y. Putra, and Y. B. Bhakti, "Developing Practicum Module of Particle Dynamics Based on Scientific Methods to Improve Students' Science Process Skills," *Sci. Educ.*, vol. 7, no. 2, p. 183, 2019, doi: 10.24235/sc.educatia.v7i2.2513.
- [21] E. R. Wulan, and H. A. Rusdiana, *Evaluasi Pembelajaran*, Bandung : Pustaka Setia, 2014.
- [22] Z. Arifin, *Evaluasi Pembelajaran: Prinsip, Teknik, Prosedur*, Bandung : PT Remaja Rosdakarya, 2012.
- [23] Asrul, R. Ananda, and Rosnita, *Evaluasi Pembelajaran*, Bandung : Citapustaka Media, 2014.
- [24] J. Widiyanto, *Evaluasi Pembelajaran: Konsep, Prinsip and Prosedur*, Madiun : UNIPMA Press, 2018.
- [25] M. Fadilah, "The Relationship Analysis Between Discriminancy Power And Validity Of Item Tests", *Proceeding of International Conference On Research, Implementation And Education Of Mathematics And Sciences*, Yogyakarta State University, 18-20 May, pp. BE-63 – BE-68, 2014.
- [26] A. Zein, M. Fadillah, and R. Novianti, "Hubungan Antara Validitas Butir, Reliabilitas, Tingkat Kesukaran Dan Daya Pembeda Soal Ujian Semester Genap Bidang Studi Biologi Kelas Xi Sma/Ma Negeri Di Kota Padang Tahun Pelajaran 2010/2011", In: *Prosiding Semirata (Seminar dan Rapat Tahunan) Bidang Ilmu MIPA*, 10-12 Mei, Universitas Lampung, pp. 39-47, 2013.
- [27] Y. Abosalem, "Assessment Techniques and Students Higher-Order Thinking Skills", *International Journal of Secondary Education*, vol. 4, no. 1, pp. 1-11, 2016, doi: 10.11648/j.ijsedu.20160401.11.

- 
- [28] M. D. Kusuma, U. Rosidin, Abdurrahman, and A. Suyatna, 2017. "The Development of Higher Order Thinking Skill (Hots) Instrument Assessment In Physics Study", *IOSR Journal of Reasearch & Method in Education (IOSR-JRME)*, vol. 7, no. 1, pp. 26-32, 2017, doi: 10.9790/7388-070103XXXX.
- [29] M. H. Bassett, "Teaching Critical Thinking without (Much) Writing: Multiple-Choice and Metacognition", *Teaching Theology & Religion*, vol. 19, no. 1, pp. 20-40, 2016, doi: 10.1111/teth.12318.
- [30] R. J. Shavelson, O. Zlatkin-Troitschanskaia, K. Beck, S. Schmidt, and J. P. Marino, "Assessment of university students' critical thinking: Next generation performance assessment", *International Journal of Testing*, vol. 19, no. 4, pp. 337-362, 2019, doi: 10.1080/15305058.2018.1543309.
- [31] C. Aloisi, and A. Callaghan, "Threats to the validity of the collegiate learning assessment (CLA +) as a measure of critical thinking skills and implications for learning gain", *Higher Education Pedagogies*, vol. 3, no. 1, pp. 137-162, 2018, doi: 10.1080/23752696.2018.1449128.