

## Karakteristik Tes Kemampuan Berpikir Kritis Siswa SMA pada Materi Momentum dan Impuls: Perbandingan Classical Theory Test (CTT) dan Model Rasch

Haridanti\*, Winny Liliawati , Yuyu Rachmat Tayubi<sup>1</sup>

Received: 5 Januari 2021 · Accepted: 13 Oktober 2021; Published Online: 28 Februari 2023

Copyright © 2023, Wahana Pendidikan Fisika



### Abstract

*The purpose of this research was to compare the results of the analysis of the critical thinking ability test instruments through classical theory test and the Rasch model in terms of the validity, reliability, difficulty level, and distinguishing power of the items. This research is quantitative descriptive. The data were obtained through a test of critical thinking skills in the form of descriptions totaling 23 items. The test was distributed to high school students in class XI and XII as many as 75 participants with 23 male and 52 female students from 21 schools via google form. Data were analyzed through classical theory tests using the Microsoft Excel program and Rasch model with Winsteps software.*

**Keywords:** Classical theory test · Critical thinking skills · Rach model

### PENDAHULUAN

Pada pembelajaran di sekolah, guru membuat tujuan dan jalur yang harus ditempuh oleh siswanya berdasarkan standar kurikulum yang ditetapkan. Guru menggunakan berbagai macam alat, media, model, metode, serta strategi sebagai penunjang pembelajaran dan tujuan yang ditetapkan dalam pembelajaran sehingga dapat melihat hasil belajar yang mampu dicapai oleh siswa. Asesmen hasil belajar siswa atau penilaian merupakan salah satu cara yang dapat digunakan untuk melihat sejauh mana pencapaian belajar siswa (Abi Hamid, 2016; Magdalena dkk., 2020).

Asesmen hasil belajar berfungsi untuk mengukur pemahaman dan penguasaan materi yang diberikan guru kepada siswa, sehingga dapat diartikan bahwa asesmen ini akan memberikan gambaran informasi berupa prestasi dan pencapaian kompetensi oleh setiap peserta didik yang diperoleh dari proses pembelajaran (Marjiastuti & Wahyuni, 2015). Asesmen memiliki peranan yang sangat penting untuk membantu siswa dalam pembelajaran, yaitu mencapai tujuan yang diciptakan oleh guru, meningkatkan motivasi belajarnya, mendorong kebiasaan belajar yang baik, bahkan memberikan umpan balik melalui identifikasi kelemahan dan kelebihan masing-masing siswa (Pisca, 2014). Oleh karena itu, asesmen yang baik mengartikan kualitas pembelajaran yang baik (Syahrul, 2010; Yus, 2018).

---

✉ Corresponding Author      Yuyu Rachmat Tayubi  
hardianti.hri@gmail.com      rachmat@upi.edu

Winny Liliawati  
winny@upi.edu

<sup>1</sup> Departemen Pendidikan Fisika, Universitas Pendidikan Indonesia, Bandung, Indonesia

Penting bagi pendidik untuk melaksanakan kegiatan asesmen yang tepat dalam mengukur kemampuan akademik peserta didik. Salah satu cara untuk mengukur kemampuan siswa ialah dengan menggunakan instrumen tes atau pun non tes (Juhairiyah, 2017). Namun, instrumen yang sering digunakan untuk mengukur kemampuan siswa di sekolah adalah instrumen tes (Magdalena dkk., 2021). Hal ini dikarenakan instrumen tes dinilai mampu menampung setiap materi yang diajarkan oleh guru kepada siswa pada pembelajaran baik dalam bentuk soal pilihan ganda maupun uraian (Muluki, 2020). Maka dari itu, soal-soal yang digunakan dalam kegiatan asesmen harus memiliki kualitas yang baik sebab hal tersebut akan menggambarkan informasi yang tepat mengenai siswa, baik terkait materi yang sudah dipahami atau belum dipahami/dikuasai oleh siswa.

Baik buruknya kualitas sebuah instrumen tes dilihat dari validitas dan reliabilitas instrumen tersebut. Validitas merupakan uji yang dilakukan untuk mengukur sejauh mana akurasi suatu tes, sedangkan reliabilitas mengukur sejauh mana hasil suatu instrumen dapat dipercaya (Syahrul, 2010; Zulpan & Rusli, 2020). Semakin tinggi nilai validitas dan reliabilitas suatu instrumen, maka akan semakin jitu data yang didapat dari suatu penelitian (Hayati & Lailatussaadah, 2016). Pernyataan tersebut didukung oleh Yusup, (2018) yang menjelaskan validitas dan reliabilitas merupakan faktor penting dalam penentuan tes yang baik atau tidak. Selain itu, terdapat aspek lain yang mendukung baiknya kualitas instrumen tes atau soal yaitu daya pembeda dan tingkat kesukaran. Daya pembeda merupakan kemampuan antara setiap butir soal untuk dapat membedakan antara siswa yang menguasai dan belum menguasai materi yang diujikan, sedangkan tingkat kesukaran merupakan kesanggupan siswa dalam menjawab soal-soal yang diujikan (Fatimah & Alfath, 2019; Hanifah, 2017). Oleh karena itu uji validitas, reliabilitas, tingkat kesukaran dan daya pembeda pada instrumen tes mutlak perlu dilakukan untuk mendapatkan instrumen tes yang berkualitas (Erfan dkk., 2020).

Instrumen tes dalam pendidikan dapat dianalisis dengan menggunakan dua pendekatan. Pendekatan pertama ialah teori tes klasik atau *Classical Theory Test* (CTT) yang paling umum dan banyak diaplikasikan dalam pendidikan hingga saat ini. Teori ini dikembangkan oleh Charles Spearman pada tahun 1904, dan digunakan untuk memprediksikan hasil suatu tes (Sumintono & Widhiarso, 2015). Pada teori tes klasik ini, aspek yang sangat menentukan baik buruknya kualitas butir soal ialah tingkat kesukaran dan daya pembeda. Namun, karakteristik butir soal yang dihasilkan oleh teori tes klasik bersifat inkonsisten (berubah) sebab bergantung pada kemampuan peserta tes. Oleh karena kekurangan tersebut, terciptalah pendekatan yang kedua yaitu pendekatan modern dengan pemodelan Rasch (Rasch model) yang dicetuskan oleh Dr. Georg Rasch seorang ahli matematika dari Denmark (Nuryanti dkk., 2018). Pemodelan Rasch memiliki pendekatan yang berbeda pada penggunaan skor atau data mentah hasil tes tersebut. Tujuan pemodelan ini ialah untuk menghasilkan suatu skala pengukuran dengan interval nilai yang sama sehingga dapat memberikan informasi yang lebih akurat baik mengenai peserta tes maupun kualitas soal tersebut. Dengan kata lain, analisis pada model Rasch akan menghasilkan informasi mengenai karakteristik butir dan siswa yang telah dibentuk menjadi metrik yang sama (Sumintono & Widhiarso, 2015)

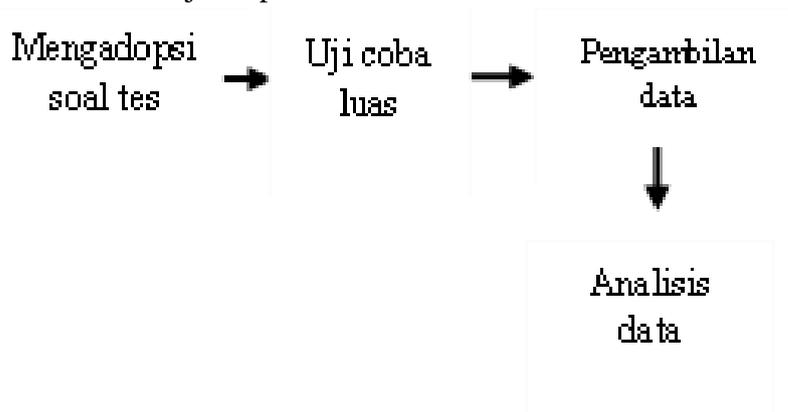
Pada penelitian ini dilakukan analisis perbandingan kualitas instrumen tes pada unsur validitas, reliabilitas, tingkat kesukaran dan daya pembeda soal melalui dua pendekatan yang telah dipaparkan di atas, yaitu teori tes klasik dan model Rasch. Instrumen tes yang digunakan pada penelitian ini adalah instrumen tes kemampuan berpikir kritis pada materi momentum dan

impuls yang diadopsi dari instrumen penelitian sebelumnya yang memiliki validitas tinggi pada 12 butir soal, 8 butir soal memiliki validitas sedang, dan 3 butir soal memiliki validitas rendah.

## METODE

### Desain Penelitian

Metode yang digunakan dalam penelitian ini yaitu metode penelitian deskriptif dan pendekatan secara kuantitatif dengan desain penelitian yaitu konstruksi dan validasi. Adapun langkah-langkah pada penelitian ini disajikan pada Gambar 1.



Gambar 1. Langkah Penelitian

Berikut ini adalah penjelasan mengenai langkah-langkah penelitian, yaitu:

#### 1. Mengadopsi soal tes

Dalam penelitian ini, tes yang digunakan diadopsi dari penelitian yang dilakukan oleh Hoerunisa (2017) yang memiliki validitas tinggi pada 12 butir soal, 8 butir soal memiliki validitas sedang, dan 3 butir soal memiliki validitas rendah.

#### 2. Uji coba luas

Tes yang telah direvisi dan di konstruksi diajukan kepada beberapa peserta didik tingkat SMA yang telah mempelajari materi momentum dan impuls. Peserta didik yang mengikuti tes sebanyak 75 orang.

#### 3. Pengambilan data

Data diperoleh dari hasil uji coba luas kepada beberapa peserta didik tingkat SMA. Data berupa jawaban siswa terhadap butir soal yang diujikan.

#### 4. Analisis data

Hasil pengolahan data, maka setiap butir soal dianalisis validitas dan reliabilitasnya dengan menggunakan CTT dan model rasch.

## Partisipan Penelitian

Partisipan penelitian ini adalah 75 orang siswa SMA dari kelas XI dan XII dengan jumlah laki-laki sebanyak 23 orang dan jumlah perempuan sebanyak 52 orang yang berasal dari 21 sekolah menengah atas (SMA) dengan mempertimbangkan bahwa partisipan telah mempelajari materi momentum dan impuls.

## Instrumen Penelitian

Instrumen penelitian yang digunakan dalam penelitian ini diadopsi dari penelitian sebelumnya yang dilakukan oleh Hoerunisa (2017) yang terdiri dari 23 butir soal yang mengukur 5 aspek keterampilan berpikir kritis yaitu: keterampilan penalaran secara verbal, keterampilan berpikir dalam hipotesis, keterampilan menganalisis dan berargumen, keterampilan menggunakan kemungkinan dan ketidakpastian, serta keterampilan dalam membuat keputusan dan pemecahan masalah. Instrumen tes tersebut sebanyak 12 butir soal memiliki validitas tinggi, 8 soal memiliki validitas sedang dan 3 soal memiliki validitas rendah

## HASIL DAN PEMBAHASAN

### Validitas

Perbandingan hasil analisis teori tes klasik dan Rasch model pada segi validitas dapat dilihat pada Tabel 1. Validitas adalah suatu ukuran yang menunjukkan tingkat-tingkat kevalidan atau kesahihan suatu instrumen. Analisis validitas butir soal melalui teori tes klasik dilakukan berdasarkan uji validitas menggunakan korelasi product moment. Sedangkan pada analisis model Rasch, uji validitas menggunakan nilai unidimensionalitas. Berdasarkan data keragaman (raw variance) yang diperoleh dari software Winstep adalah 60,9% tidak jauh beda dengan nilai ekspektasinya yaitu 60,7%. Hal ini menunjukkan persyaratan minimum batas unidimensi rasch terpenuhi yaitu diatas 40% (Linacre, 2011). Hal tersebut menunjukkan bahwa instrumen tes tersebut mampu mengukur keragaman kemampuan berpikir kritis siswa.

Berdasarkan hasil analisis kedua pendekatan, pendekatan klasik dan model rasch menyatakan bahwa instrumen tes memiliki kategori yang valid, akan tetapi memiliki hasil analisis yang berbeda. Untuk mendapatkan nilai validitas, pendekatan klasik lebih dapat memberikan gambaran kualitas setiap butir soal. Karena dalam klasik setiap butir soal dianalisis. Sedangkan dalam pendekatan model rasch, nilai validitas hanya digambarkan oleh besarnya nilai unidimensionalitasnya.

**Tabel 1.** Perbandingan Hasil Analisis Validitas Butir Soal Melalui Pendekatan Teori Tes Klasik (CTT) dan Model Rasch

No Soal	Teori tes klasik	Model Rasch
1	Valid	
2	Valid	
3	Valid	
4	Valid	
5	Valid	
6	Valid	
7	Valid	
8	Valid	
9	Valid	
10	Valid	Valid
11	Valid	
12	Valid	
13	Valid	
14	Valid	
15	Valid	
16	Valid	
17	Valid	
18	Valid	
19	Valid	

No Soal	Teori tes klasik	Model Rasch
20	Valid	
21	Valid	
22	Valid	
23	Valid	

**Reliabilitas**

**Tabel 2.** Perbandingan Hasil Reliabilitas Instrumen melalui Pendekatan *Classical Theory Test* dan Model Rasch

Classical Theory Test		Model Rasch					
Alpha Cronbach	Kategori	Person Reliability	Kategori	Item Reliability	Kategori	Alpha Cronbach	Kategori
1,00	Sangat Baik	0,92	Bagus Sekali	0,96	Istimewa	1,00	Istimewa

Menurut (Sugiyono, 2018) instrumen yang reliabel adalah instrumen yang bila digunakan beberapa kali untuk mengukur objek yang sama, akan menghasilkan data yang sama. Pada analisis teori tes klasik, menguji reliabilitas menggunakan nilai *Alpha Cronbach*, sedangkan pada analisis model rasch nilai reliabilitas didapat dari Tabel 2 pada aplikasi Winstep. Tabel 2 di atas memberikan informasi mengenai hasil analisis instrumen tes berpikir kritis pada unsur reliabilitas melalui pendekatan *Classical Theory Test*. Nilai reliabilitas yang diperoleh yaitu sebesar 1,00 dengan kategori yang sangat baik. Sedangkan analisis melalui model rasch diperoleh nilai *Item Reliabilitas* sebesar 0,96 dan *Person Reliability* sebesar 0,92 yang keduanya memiliki kategori yang sangat baik. Selain itu nilai *Alpha Cronbach* yang diperoleh dari analisis model rasch yaitu sebesar 1,00 dengan kategori sangat baik.

**Tingkat Kesukaran**

**Tabel 3.** Perbandingan Hasil Analisis Tingkat Kesukaran Butir Soal Melalui Pendekatan Teori Tes Klasik dan Model Rasch

No Soal	Teori tes klasik	Model Rasch
1	Mudah	Sangat Mudah
2	Sedang	Mudah
3	Sedang	Sukar
4	Mudah	Mudah
5	Sedang	Mudah
6	Sedang	Sukar
7	Sedang	Sukar
8	Sedang	Mudah
9	Mudah	Sangat Mudah
10	Sedang	Mudah
11	Sedang	Mudah
12	Sedang	Sangat Sukar
13	Sedang	Mudah
14	Sedang	Mudah
15	Sedang	Sangat Mudah
16	Sedang	Mudah
17	Sedang	Mudah
18	Sedang	Sukar
19	Sedang	Mudah
20	Sedang	Sukar
21	Sedang	Sukar
22	Sedang	Sangat Sukar



No Soal	Teori tes klasik	Model Rasch
23	Sedang	Sangat Sukar

Indeks kesukaran soal adalah bilangan yang menunjukkan sukar atau mudahnya suatu soal (Arikunto, 2010). Berdasarkan Tabel 3 menurut teori tes klasik terdapat 20 soal memiliki nilai tingkat kesukaran dengan kategori sedang dan terdapat tiga butir soal yang memiliki kategori mudah. Sedangkan menurut model rasch, 23 butir tes keterampilan berpikir kritis terdapat tiga butir soal dengan kategori sangat sukar, enam butir soal dengan kategori sukar, 11 butir soal dengan kategori mudah dan tiga butir soal dengan kategori sangat mudah.

### Daya Pembeda

Daya pembeda suatu soal merupakan kemampuan soal untuk membedakan antara siswa yang memiliki tingkat kemampuan yang tinggi dengan siswa yang memiliki kemampuan yang rendah. Menurut (Novytsari, 2014) indeks daya pembeda sudah dapat dinyatakan layak apabila lebih dari 0,20 (dengan kriteria cukup).

Hasil analisis daya pembeda soal melalui pendekatan teori tes klasik menunjukkan bahwa 21 soal memiliki nilai daya pembeda dengan kategori cukup dan terdapat dua soal yang memiliki kategori jelek yaitu nomor 12 dan nomor 23. Artinya 23 butir soal yang telah dianalisis berdasarkan indeks daya bedanya melalui pendekatan teori tes klasik hanya ada dua soal yang merupakan soal yang tidak layak untuk digunakan.

Berbeda dengan pendekatan teori tes klasik, pada analisis melalui model Rasch, untuk dapat membedakan kemampuan siswa yang tinggi dan rendah digunakan analisis pada tingkat abilitas individu. Selain itu dapat pula dilihat dengan cara mengidentifikasi kelompok responden berdasarkan indeks separasi responden. Menurut (Sumintono & Widhiarso, 2015) semakin besar nilai separasi maka kualitas instrumen dalam hal keseluruhan responden dan butir makin bagus, karena dapat mengidentifikasi kelompok responden dan butir. Persamaan lain untuk melihat pengelompokkan secara lebih teliti digunakan persamaan pemisahan strata (H):

$$H = \frac{[(4 \times SEPARATION) + 1]}{3} \quad (1)$$

Dari nilai separasi butir sebesar 4,65 maka  $H = 6,41$  dibulatkan menjadi 6. Artinya terdapat enam kelompok butir soal yang dapat diidentifikasi dan untuk responden yang diperoleh sebesar 3,37, maka  $H = 4,83$  dibulatkan menjadi 5, menunjukkan bahwa kelompok responden dapat dibedakan menjadi lima kelompok yang menyatakan kemampuan peserta didik.

### SIMPULAN

Berdasarkan pembahasan diatas, dapat ditarik kesimpulan bahwa kedua analisis mampu mengkarakterisasi instrumen tes. Kedua analisis menghasilkan hasil yang sama pada nilai validitas dan reliabilitas dengan hasil yang sangat baik. Untuk nilai tingkat kesukaran, pada analisis CTT menghasilkan dua kategori sedangkan pada model rasch menghasilkan empat kategori. Pada analisis CTT, nilai daya pembeda menghasilkan dua kategori, sedangkan pada model rasch menghasilkan lima kategori. Dilihat dari hasil analisis, model rasch mampu menganalisis lebih baik daripada CTT dalam mengukur reliabilitas, tingkat kesukaran dan daya

pembeda. Selain empat hal di atas, model rasch mampu mengukur kemampuan abilitas peserta didik. Berdasarkan penelitian yang telah dilakukan, penelitian lain dapat dilakukan pada materi yang berbeda dengan proses pengambilan data berulang. Hal tersebut dapat menyempurnakan data yang didapat.

## REFERENCES

- Abi Hamid, M. (2016). Pengembangan instrumen penilaian hasil belajar siswa berbasis tik pada pembelajaran dasar listrik elektronika. *VOLT: Jurnal Ilmiah Pendidikan Teknik Elektro*, 1(1), 37–46.
- Arikunto, S. (2010). *Prosedur penelitian suatu pendekatan praktek*. Dalam Jakarta: Rineka Cipta.
- Erfan, M., Maulyda, M. A., Hidayati, V. R., Astria, F. P., & Ratu, T. (2020). Analisis kualitas soal kemampuan membedakan rangkaian seri dan paralel melalui teori tes klasik dan model Rasch. *Indonesian Journal of Educational Research and Review*, 11–19.
- Fatimah, L. U., & Alfath, K. (2019). Analisis kesukaran soal, daya pembeda dan fungsi distraktor. *AL-MANAR: Jurnal Komunikasi dan Pendidikan Islam*, 8(2), 37–64.
- Hanifah, N. (2017). Perbandingan Tingkat Kesukaran, Daya Pembeda Butir Soal Dan Reliabilitas Tes Bentuk Pilihan Ganda Biasa Dan Pilihan Ganda Asosiasi Mata Pelajaran Ekonomi. *Sosio e-KONS*, 6(1), 41–55.
- Hayati, S., & Lailatussaadah, L. (2016). Validitas dan reliabilitas instrumen pengetahuan pembelajaran aktif, kreatif dan menyenangkan (pakem) menggunakan model Rasch. *Jurnal Ilmiah Didaktika: Media Ilmiah Pendidikan dan Pengajaran*, 16(2), 169–179.
- Hoerunisa, A. (2017). *Karakterisasi Tes Keterampilan Berpikir Kritis Materi Momentum dan Impuls menggunakan Item Response Theory*, Universitas Pendidikan Indonesia.
- Juhairiyah, J. (2017). Assesmen Konten Isi Bidang Studi. *PEDAGOGIK: Jurnal Pendidikan*, 4(1).
- Linacre, J. M. (2011). *A User's guide to WINSTEPS Ministeps; Rasch-model Computer Program*. Program Manual 3.73.
- Magdalena, I., Fauzi, H. N., & Putri, R. (2020). Pentingnya evaluasi dalam pembelajaran dan akibat memanipulasinya. *Bintang*, 2(2), 244–257.
- Magdalena, I., Mahromiyati, M., & Nurkamilah, S. (2021). Analisis Instrumen Tes Sebagai Alat Evaluasi Pada Mata Pelajaran SBdP Siswa Kelas II SDN Duri Kosambi 06 Pagi. *NUSANTARA*, 3(2), 276–287.
- Marjiastuti, K., & Wahyuni, S. (2015). *Analisis kemampuan peserta didik dengan model Rasch*. Seminar Nasional Evaluasi Pendidikan II.
- Muluki, A. (2020). Analisis kualitas butir tes semester ganjil mata pelajaran IPA Kelas IV MI Radhiatul Adawiyah. *Jurnal Ilmiah Sekolah Dasar*, 4(1), 86–96.
- Novytasari, Y. P. (2014). *Analisis Butir Soal Pilihan Ganda Menggunakan Teori Pengukuran Klasik Pada Ulangan Umum Akhir Semester Genap Bahasa Prancis SMA Negeri 9 Yogyakarta Tahun Ajaran 2013/2014*. Universitas Negeri Yogyakarta.
- Nuryanti, S., Masykuri, M., & Susilowati, E. (2018). Analisis Iteman dan model Rasch pada pengembangan instrumen kemampuan berpikir kritis peserta didik sekolah menengah kejuruan. *Jurnal Inovasi Pendidikan IPA*, 4(2), 224–233.
- Pisca, C. C. (2014). *Analisis Perbandingan Kualitas Butir Soal Ujian Sekolah Bahasa Prancis SMAN 10 Yogyakarta Tahun Ajaran 2013/2014 Dilihat dari Paradigma Klasik dan Modern (PhD Thesis)*. Universitas Negeri Yogyakarta.
- Sugiyono, S. (2018). *Metode Penelitian Pendidikan Pendekatan Kualitatif, Kuantitatif dan R & D*. Alfabeta, Bandung.
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan rasch pada assessment pendidikan*. Trim komunikasi.

- Syahrul, S. (2010a). Pengembangan Model Asesmen Kompetensi Siswa SMK dalam Konteks Pembelajaran Berbasis Kerja di Industri. *Jurnal Penelitian dan Evaluasi Pendidikan*, 14(2), 246–268.
- Yus, A. (2018). *Konsep Asesmen, Evaluasi dan Kaitannya dengan Pembelajaran*. Pustaka Univeristas Terbuka
- Yusup, F. (2018). Uji validitas dan reliabilitas instrumen penelitian kuantitatif. *Tarbiyah: Jurnal Ilmiah Kependidikan*, 7(1).
- Zulpan, Z., & Rusli, A. (2020). Validitas dan Reliabilitas Instrumen Penilaian Membaca Short Functional Text pada Siswa SMP Kelas VIII. *Jurnal Pendidikan Guru*, 1(1).

