

## PENGEMBANGAN TES KETERAMPILAN BERPIKIR KRITIS BERDASARKAN ANALISIS TEORI RESPON BUTIR

Desi Ratna Ningsih\*), Taufik Ramlan Ramalis, Unang Purwana

Universitas Pendidikan Indonesia, Jl. Dr. Setiabudhi 229, Bandung 40154, Jawa Barat

\* Email : desiratna1609@gmail.com

### ABSTRAK

Keterampilan berpikir kritis merupakan salah satu keterampilan yang harus dimiliki oleh siswa sebagai standar kompetensi lulusan. Namun, keterampilan ini masih jarang dinilai karena belum banyaknya instrumen untuk mengukur keterampilan tersebut. Oleh karena itu, dilakukan penelitian untuk mengembangkan instrumen tes keterampilan berpikir kritis yang berbentuk pilihan ganda dengan menggunakan metode *Research and Development (R&D)* yang melibatkan tiga ahli. Hasil validasi isi oleh tiga ahli yang dianalisis menggunakan Aiken V menunjukkan bahwa instrumen tes ini valid dan layak diuji coba. Instrumen tes diuji coba kepada siswa SMA dan hasilnya dianalisis menggunakan teori respon butir dengan bantuan program *eirt* versi 2.0.0. Hasil analisis menunjukkan bahwa model 3PL merupakan model yang sesuai untuk mengkarakteristik instrumen tes keterampilan berpikir kritis. Instrumen tes ini memiliki fungsi informasi sebesar 30,14 dengan *Standard Error of Measurement (SEM)* sebesar -0.83, daya pembeda, tingkat kesukaran, dan tebakan semu dengan kategori "baik".

Kata Kunci: keterampilan berpikir kritis, teori respon butir, model parameter logistik

### ABSTRACT

Critical thinking skill is one of the skills that must be possessed by students as the graduate competency standards. However, this skill is rarely rated as not many instruments to measure the skill. This research was conducted to develop critical thinking skills test instruments in the form of multiple choice with Research and Development (R&D) as the method that involving three experts. The results of the content validation by three experts who analyzed using Aiken V indicates that the test instrument is valid and worthy of being tested. The instrument tests to high school students tested and the results were analyzed using item response theory program of *eirt* version 2.0.0. The results show that the model 3PL is an appropriate model to characterize critical thinking skills test instruments. The test instrument has the function of information at 30.14 with the Standard Error of Measurement (SEM) of -0.83, distinguishing, level of difficulty, and guess pseudo with the "good" category.

Keywords: critical thinking skills; item response theory; parameter logistic models

### PENDAHULUAN

*Departement of Defense Education Activity (DoDEA)* merumuskan ada tujuh keterampilan yang harus dicapai dalam *21st Century Teaching, Learning, and Leading (21st CTLL)* sebagai bekal kesuksesan siswa di abad 21 yaitu, *critical thinking and problem solving, collaboration across networks, agility and adaptability, initiative and entrepreneurship, effective oral and written communication, accessing and analyzing information, dan curiosity and imagination* [1]. Salah satu keterampilan yang harus dimiliki oleh siswa menurut DoDEA yaitu keterampilan berpikir

kritis. Permendikbud Nomor 20 Tahun 2016 tentang Standar Kompetensi Lulusan Pendidikan Dasar dan Menengah, menyatakan bahwa pada dimensi keterampilan siswa harus memiliki keterampilan berpikir dan bertindak kreatif, produktif, kritis, mandiri, kolaboratif, dan komunikatif melalui pendekatan ilmiah sebagai pengembangan dari yang dipelajari di satuan pendidikan dan sumber lain secara mandiri. Berdasarkan uraian tersebut, maka keterampilan berpikir kritis merupakan kompetensi yang penting untuk dilatihkan kepada siswa [2].

Keterampilan berpikir kritis merupakan suatu kemampuan yang dapat menciptakan para pemikir yang tangguh dan pemecah masalah yang handal [3]. Terdapat enam aspek yang dapat digunakan untuk mengetahui keterampilan berpikir kritis seseorang, yaitu eksplanasi, interpretasi, inferensi, analisis, evaluasi, dan regulasi diri [1,3]. Interpretasi (*interpretation*), meliputi kategorisasi, pemecahan kode, dan klarifikasi makna. Analisis (*analysis*), meliputi menguji ide atau gagasan, mendeteksi argumen, dan menganalisis argumen. Inferensi (*inference*), meliputi mengumpulkan informasi, dugaan alternatif, dan membuat kesimpulan. Evaluasi (*evaluation*), meliputi penilaian klaim dan penilaian argumen. Eksplanasi (*explanation*), meliputi menyatakan hasil, membenarkan suatu prosedur, dan menyajikan argumen. Regulasi diri (*self-regulation*), meliputi pemeriksaan dan pengkoreksian diri.

Keterampilan berpikir kritis bukanlah keterampilan yang dibawa oleh seseorang sejak ia lahir, sehingga keterampilan ini dapat diterapkan, dilatih, dan dikembangkan melalui proses pembelajaran dan evaluasi pembelajaran. Dalam proses pembelajaran, guru berperan sebagai mediator dan fasilitator. Sehingga dalam pelaksanaannya, guru dapat mendesain dan menerapkan pendekatan, model, dan metode pembelajaran yang dapat memfasilitasi siswa dalam meningkatkan keterampilan berpikir kritis. Sementara itu dalam evaluasi pembelajaran, guru dapat memberikan latihan kepada siswa melalui pemberian masalah dalam bentuk soal yang bervariasi dan juga guru dapat mengembangkan instrumen yang mampu mengukur keterampilan berpikir kritis siswa.

Proses pengembangan instrumen dapat dilakukan dalam bentuk tes dan nontes. Ennis, Millman, & Tomko mengembangkan *Cornell Critical Thinking Test-Level Z (CCTT)* dalam bentuk pilihan ganda [4]. Ennis & Wier mengembangkan *Ennis-Wier Critical Thinking Essay Test* dalam bentuk essay (*open-ended*) [5]. Facione mengembangkan *California Critical Thinking Skills Test (CCTST)* dalam bentuk pilihan ganda [6]. Halpern mengembangkan instrumen dalam dua bentuk yaitu pilihan ganda dan uraian [7]. Tiruneh, et al. mengembangkan instrumen *Critical Thinking Test in Electricity and Magnetism (CTEM)* dalam bentuk uraian [8]. Pengembangan instrumen keterampilan berpikir kritis dalam mata pelajaran tertentu misalnya fisika, bisa

melalui aspek kognitif. Sehingga pengembangan instrumen pada penelitian ini berupa tes dalam bentuk pilihan ganda dengan materi fisika yaitu getaran harmonis.

Untuk memperoleh instrumen tes yang valid dalam mengukur keterampilan berpikir kritis yang dimiliki siswa, dapat dilakukan validasi isi. Validasi isi instrumen adalah analisis rasional yang dikemukakan oleh ahli untuk mengetahui keterwakilan instrumen dengan kemampuan yang hendak diukur [9]. Validasi isi mencakup aspek materi, konstruk, dan Bahasa yang terdapat dalam instrumen. Untuk mengetahui kesepakatan ahli, maka dapat digunakan indeks validitas yang diusulkan oleh Aiken yang dirumuskan sebagai berikut.

$$V = \frac{\sum s}{n(c - 1)}$$

dengan  $V$  adalah indeks kesepakatan ahli mengenai validitas isi butir;  $s$  adalah skor yang ditetapkan ahli dikurangi skor terendah dalam kategori yang digunakan ( $s = r - l_0$ , dengan  $r$  adalah skor kategori pilihan ahli dan  $l_0$  adalah skor terendah dalam kategori penilaian);  $n$  adalah banyaknya ahli; dan  $c$  adalah banyaknya kategori yang dapat dipilih oleh ahli. Indeks Aiken  $V$  nilainya berkisar antara 0 sampai 1 dengan kategori tingkat validitas tinggi jika  $V > 0,8$ ; tingkat validitas sedang jika  $0,4 < V \leq 0,8$ ; dan tingkat validitas kurang jika  $V \leq 0,4$ .

Untuk menghasilkan instrumen tes yang baik, harus diketahui pula karakteristik instrumen tes yang lain, yaitu daya pembeda, tingkat kesukaran, tebakan semu, dan reliabilitas. Dalam penelitian ini, karakteristik instrumen tes dapat diketahui dengan menggunakan analisis teori respon butir dengan menggunakan bantuan program *eirt* versi 2.0.0 dan model tiga parameter logistik (3PL). Secara matematis model tiga parameter logistik dirumuskan sebagai berikut [10].

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{a_i(\theta - b_i)}}{1 + e^{a_i(\theta - b_i)}}$$

dengan

$P_i(\theta)$  : peluang bahwa peserta tes dengan kemampuan menjawab butir soal ke- $i$  dengan benar.

$a_i$  : parameter daya pembeda soal butir ke- $i$ .

$b_i$  : parameter tingkat kesukaran, yaitu satu titik pada skala *ability* di mana kemungkinan untuk menjawab benar sebesar 0,5.

$c_i$  : peluang tebakan benar butir ke- $i$ .

$\theta$  : parameter kemampuan peserta tes.

$e$  : bilangan natural yang nilainya mendekati 2,718.

Pada teori respon butir juga dikenal istilah fungsi informasi yang terdiri dari butir tes dan tes secara keseluruhan. Fungsi informasi butir (*item information functions*) merupakan suatu metode untuk menggambarkan kekuatan suatu butir pada perangkat tes, pemilihan butir tes dan perbandingan beberapa perangkat tes. Dengan fungsi informasi butir, maka dapat diketahui butir mana yang cocok dengan model teori respons butir sehingga dapat membantu dalam menyeleksi butir tes [10]. Secara matematis, fungsi informasi butir ditulis sebagai berikut.

$$I_i(\theta) = \frac{[P'_i(\theta)]^2}{P_i(\theta)Q_i(\theta)}$$

dengan

$i$  : 1,2,3,...,n

$I_i(\theta)$  : fungsi informasi butir ke-i

$P_i(\theta)$  : peluang peserta dengan kemampuan  $\theta$  menjawab benar butir i

$P'_i(\theta)$  : turunan pertama fungsi  $P_i(\theta)$  terhadap  $\theta$

$Q_i(\theta)$  : peluang peserta dengan kemampuan  $\theta$  menjawab salah butir i.

Fungsi informasi pada setiap butir tes pada perangkat tes menunjukkan fungsi informasi tes tersebut secara keseluruhan sehingga jika  $FI(\theta)$  menunjukkan informasi tes maka nilainya sama dengan jumlah dari fungsi informasi penyusun tes tersebut [10,11]. Secara matematis, fungsi informasi keseluruhan suatu butir tes sebagai berikut.

$$FI(\theta) = \sum_{j=1}^n I_j$$

Nilai indeks parameter butir dan kemampuan peserta merupakan hasil estimasi sehingga kebenarannya bersifat probabilitas dan tidak terlepas dari kesalahan dalam pengukuran. Pada teori respons butir, kesalahan dalam pengukuran dikenal sebagai kesalahan penaksiran standar (*Standard Error of Measurement, SEM*) yang nilainya berkaitan dengan fungsi informasi. Hubungan SEM dan fungsi informasi adalah berbanding terbalik kuadrat sehingga semakin besar fungsi informasi maka SEM semakin kecil atau sebaliknya [11]. Hubungan matematis nilai estimasi SEM dan fungsi informasi sebagai berikut.

$$SEM(\theta) = \frac{1}{\sqrt{I(\theta)}}$$

## METODE

*Research and Development* (R&D) adalah metode penelitian yang digunakan. Langkah-langkah penelitian R&D yaitu, potensi dan masalah, mengumpulkan informasi, desain produk, validasi desain, perbaikan desain, uji coba produk, revisi produk, uji coba pemakaian, revisi produk, dan pembuatan produk masal [12]. Namun, dalam penelitian ini terjadi modifikasi langkah-langkah penelitian sehingga hanya enam dari sepuluh langkah penelitian metode R&D yang dilakukan oleh peneliti. Produk instrumen tes keterampilan berpikir kritis yang dihasilkan membahas mengenai materi getaran harmonis untuk siswa SMA. Oleh karena itu, populasi dalam penelitian yaitu siswa SMA yang telah mempelajari materi fisika tersebut. Namun, yang dijadikan sampel hanya 75 orang siswa SMA di salah satu SMA Negeri di Kota Bandung. Penarikan sampel menggunakan teknik *nonprobability sampling* jenis *purposive sampling*.

## HASIL DAN PEMBAHASAN

Berdasarkan teori yang telah dikemukakan, didesain instrumen tes keterampilan berpikir kritis yang bertujuan untuk mengukur keterampilan berpikir kritis yang dimiliki oleh siswa. Instrumen tes yang dikembangkan berbentuk pilihan ganda karena instrumen bentuk tersebut sudah banyak digunakan yang berarti mampu untuk mengukur keterampilan berpikir kritis yang dimiliki oleh siswa. Instrumen tes yang dikembangkan mengacu pada aspek yang dikemukakan oleh Facione, namun tidak semua aspek digunakan [13]. Instrumen tes ini dikembangkan sebanyak 20 butir soal dengan lima pilihan jawaban dan masing-masing aspek diwakili sebanyak empat butir soal.

Untuk mengetahui kesesuaian antara instrumen tes yang didesain dan keterampilan berpikir kritis yang akan diukur, maka dilakukan validasi isi dengan melibatkan tiga ahli. Setiap ahli memberikan penilaian dan saran terhadap instrumen yang telah didesain pada lembar *judgement* validasi isi. Untuk mengetahui hasil kesepakatan tiga ahli digunakan analisis indeks Aiken V. Hasil kesepakatan ahli berdasarkan indeks Aiken V disajikan dalam tabel 1 berikut ini.

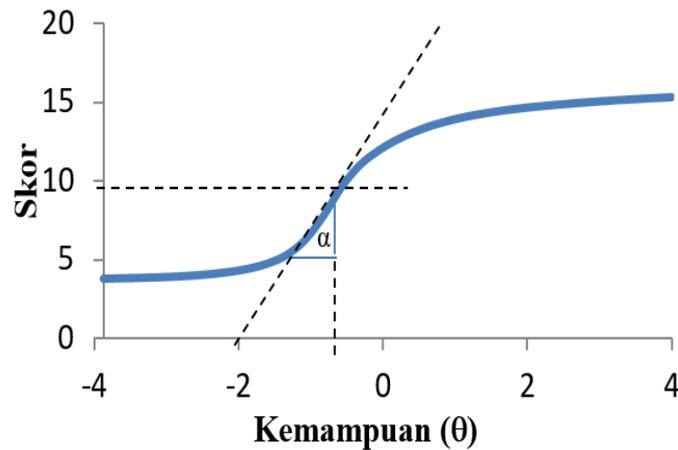
**Tabel 1. Hasil Analisis Indeks Aiken V pada Validasi Isi Tes Keterampilan Berpikir Kritis**

No	Aspek	Indikator	Indeks Aiken V	Interpretasi	Rerata dan Interpretasi
1.	Materi	Kesesuaian dengan KI dan KD	0.83	Validitas tinggi	0.78 Validitas sedang
		Kesesuaian dengan indikator soal	0.83	Validitas tinggi	
		Kesesuaian dengan aspek keterampilan berpikir kritis	0.67	Validitas sedang	
2.	Konstruk	Pilihan jawaban homogen dan logis.	0.67	Validitas sedang	0.89 Validitas tinggi
		Hanya ada satu kunci jawaban.	1.00	Validitas tinggi	
		Pokok soal dirumuskan dengan singkat, jelas, dan tegas.	0.67	Validitas sedang	
		Pokok soal bebas dari pernyataan yang bermakna ganda.	1.00	Validitas tinggi	
		Gambar, grafik, tabel atau sejenisnya jelas dan berfungsi.	1.00	Validitas tinggi	
		Pilihan jawaban yang berbentuk angka disusun berdasarkan urutan besar kecilnya angka.	1.00	Validitas tinggi	
3.	Bahasa	Butir soal menggunakan bahasa Indonesia yang baku dan sesuai dengan EYD.	1.00	Validitas tinggi	1.00 Validitas tinggi
		Rumusan butir soal menggunakan bahasa yang komunikatif.	1.00	Validitas tinggi	
<b>Rerata</b>					<b>0.89 Validitas tinggi</b>

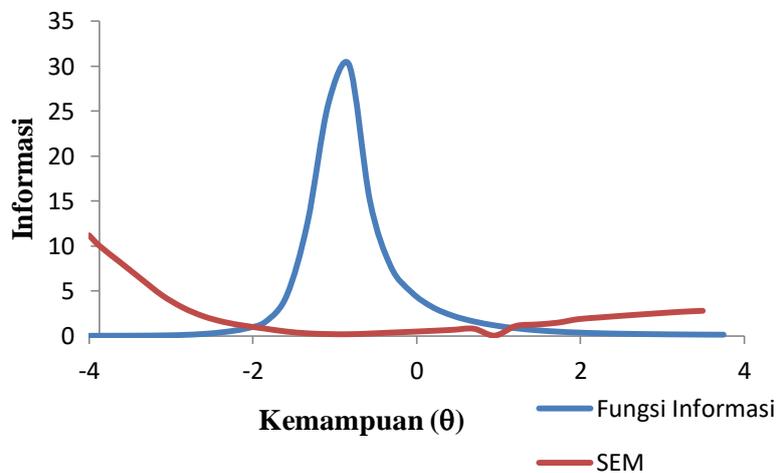
Berdasarkan tabel 1 diketahui bahwa indeks Aiken V dari instrumen tes keterampilan berpikir kritis sebesar 0,89 dengan kategori validitas tinggi. Artinya, ketiga ahli menyatakan bahwa instrumen tes keterampilan berpikir kritis yang didesain oleh peneliti layak diuji coba, sehingga dapat diketahui karakteristiknya.

Data hasil uji coba diperoleh dari jawaban siswa yang dikoreksi dengan menggunakan penilaian data dikotomi yang memiliki nilai 1 apabila jawaban benar dan 0 apabila jawaban salah sehingga diperoleh data hasil uji coba

berupa angka. Angka tersebut kemudian dianalisis menggunakan teori respon butir dengan bantuan program *eirt* versi 2.0.0., sehingga diketahui karakteristik butir tesnya. Hasil analisis menunjukkan bahwa model 3PL merupakan model yang sesuai untuk mengkararakteristik tes tersebut. Model 3PL menitikberatkan pada daya pembeda (a), tingkat kesukaran (b), dan tebakan semu (c). Nilai daya pembeda, tingkat kesukaran, dan tebakan semu secara keseluruhan dapat diketahui dari kurva karakteristik total, seperti yang disajikan pada gambar 1 berikut ini.



Gambar 1. Kurva Karakteristik Total Model 3PL



Gambar 2. Fungsi Informasi dan SEM Model 3PL

Berdasarkan gambar 1, diketahui bahwa nilai daya pembeda sebesar 1,13 dengan kategori baik, nilai tingkat kesukaran sebesar -0,83 dengan kategori baik, dan nilai tebakan semu sebesar 0,20 dengan kategori baik. Karakteristik lainnya yaitu estimasi reliabilitas yang menunjukkan konsistensi instrumen tes dalam melakukan pengukuran kemampuan berpikir kritis. Estimasi reliabilitas diperoleh dari titik perpotongan fungsi informasi dan SEM, seperti gambar 2 berikut ini.

Berdasarkan gambar 2, diketahui bahwa titik perpotongan antara fungsi informasi total dan SEM berada pada rentang kemampuan -1,84 sampai 1,21 dengan puncak informasi total sebesar 30,14 dan SEM sebesar -0,83. Hal tersebut menunjukkan bahwa instrumen tes keterampilan berpikir kritis memiliki estimasi reliabilitas pada kategori baik serta dapat memberikan informasi kemampuan yang dimiliki oleh siswa secara bervariasi.

## KESIMPULAN

Tes keterampilan berpikir kritis hasil pengembangan layak diuji coba kepada siswa SMA yang telah mempelajari materi getaran harmonis. Berdasarkan hasil uji coba diketahui bahwa semua butir tes memiliki karakteristik yang berkategori baik.

## UCAPAN TERIMA KASIH

Puji dan syukur senantiasa penulis ucapkan kehadiran Allah SWT., atas rahmat dan karunia-Nya sehingga penulis mampu menyelesaikan artikel ini. Penulis juga mengucapkan terima kasih kepada Bapak Dr. Taufik Ramlan Ramalis, M.Si., dan Bapak Drs. Unang Purwana, M.Pd., yang telah membimbing penulis selama penelitian hingga penyelesaian artikel ini.

## REFERENSI

- [1] Facione, P. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (The Delphi Report).
- [2] No, P. (21). Tahun 2016. *Standar Isi pendidikan Dasar dan Menengah yang memuat tentang Tingkat Kompetensi dan Kompetensi Inti Sesuai dengan Jenjang dan Jenis Pendidikan tertentu*.
- [3] Facione, P. A. (1998). Critical thinking: What it is and why it counts. Retrieved June, 9, 2004.
- [4] Ennis, R. H., Millman, J., & Tomko, T. N. (1985). *Cornell critical thinking tests level X & level Z: Manual*. Pacific Grove, CA: Midwest Publications.
- [5] Ennis, R. H., & Weir, E. E. (1985). *The Ennis-Weir critical thinking essay test: An instrument for teaching and testing*. Midwest Publications.
- [6] Facione, P. A. (1990). The California Critical Thinking Skills Test--College Level. Technical Report# 1. Experimental Validation and Content Validity.
- [7] Halpern, D. F. (2010). Halpern critical thinking assessment. *SCHUHFRIED (Vienna Test System): Moedling, Austria*. [Http://www. Schuhfried. com/vienna-Test-System-vts/all-Tests-fromaz/test/hcta-Halpern-Critical-Thinking-Assessment](http://www.Schuhfried.com/vienna-Test-System-vts/all-Tests-fromaz/test/hcta-Halpern-Critical-Thinking-Assessment), 1.
- [8] Tiruneh, D. T., De Cock, M., Weldeslassie, A. G., Elen, J., & Janssen, R. (2017). Measuring critical thinking in physics: Development and validation of a critical thinking test in electricity and magnetism. *International Journal of Science and Mathematics Education*, 15(4), 663-682.
- [9] Retnawati, H. (2016). Analisis kuantitatif instrumen penelitian. Yogyakarta: *Parama publising*.
- [10] Retnawati, H. (2014). Teori respon butir dan penerapannya. Yogyakarta: *Nuha Medika*.
- [11] Naga, D. S. (1992). Pengantar teori skor pada pengukuran pendidikan. Jakarta: *Gunadarma*.
- [12] Sugiyono. (2008). *Metode penelitian pendidikan:(pendekatan kuantitatif, kualitatif dan R & D)*. Alfabeta.
- [13] Facione, P. A., & Facione, N. C. (2013). Critical thinking for life: Valuing, measuring, and training critical thinking in all its forms. *Inquiry: Critical Thinking Across the Disciplines*, 28(1), 5-25.