

Instrumen Tes untuk Mengukur Keterampilan Berpikir Kritis pada Materi Statistika

Hanief Abdur Rahman^{1✉}, Bachtiar Syaiful Bachri², Irena Y. Maureen³

^{1, 2, 3} Teknologi Pendidikan, Fakultas Ilmu Pendidikan, Universitas Negeri Surabaya, Jl. Lidah Wetan, Surabaya, Indonesia
abdurrahman.hanief@gmail.com

Abstract

To raise the standard of instruction, teachers must be aware of their pupils' cognitive abilities. The purpose of this project is to create a test that will assess students' critical thinking abilities when it comes to statistics content. The development process is conducted in three stages: define, design, and develop. With i-CVI values > 0.83 and s-CVI = 0.95, the validator has deemed the test instrument generated to be legitimate, meaning it satisfies the content validity requirements. 38 students from class XI at SMA Kemala Bhayangkari 3 Porong were then used to test the instrument. The critical thinking test instrument was determined to be reliable with a Cronbach's Alpha value of 0.656 and met the valid criterion with a computed r value > 0.320 based on the test results. Additionally, the test instrument produced has a moderate level of difficulty, with a difficulty index value in the range of 0.3 to 0.8, and good discrimination, with a discrimination index value of > 0.35 . The study's test subjects were constrained, though. Thus, additional research with participants with varying educational backgrounds is required in order to determine the applicability and efficacy of the test instruments produced in assessing the critical thinking abilities of students from a wider range of backgrounds.

Keywords: Critical Thinking Test, Statistics, Validity, Reliability.

Abstrak

Mengetahui tingkat keterampilan berpikir peserta didik menjadi hal penting bagi seorang pendidik guna meningkatkan kualitas pendidikan. Penelitian ini bertujuan mengembangkan instrumen tes yang bisa mengukur keterampilan berpikir kritis peserta didik pada materi statistika. Ada tiga tahapan pengembangan yang dilakukan, yaitu *define*, *design*, dan *develop*. Instrumen tes yang dikembangkan telah dikatakan valid validator dengan nilai i-CVI $> 0,83$ dan s-CVI = 0,95 sehingga memenuhi kriteria validitas isi. Selanjutnya, instrumen tersebut diujicobakan kepada 38 peserta didik kelas XI SMA Kemala Bhayangkari 3 Porong untuk menguji validitas konstruk, reliabilitas, daya beda, serta tingkat kesukaran soal. Dari hasil ujicoba tersebut didapat instrumen tes berpikir kritis sudah memenuhi kriteria valid dengan nilai $r_{hitung} > 0,320$ serta reliabel dengan nilai Cronbach's Alpha senilai 0,656. Instrumen tes yang dikembangkan juga memiliki daya beda yang baik dengan nilai indeks daya beda $> 0,35$ serta tingkat kesukaran sedang dengan nilai indeks tingkat kesukaran yang ada dalam interval 0,3 – 0,8. Namun, subjek uji penelitian ini dibatasi. Untuk menilai kegunaan dan kemampuan instrumen tes yang dibuat untuk mengukur kemampuan berpikir kritis siswa dari berbagai latar belakang yang lebih luas, diperlukan lebih banyak penelitian dengan peserta dari latar belakang pendidikan yang berbeda-beda.

Kata kunci: Tes Berpikir Kritis, Statistika, Validitas, Reliabilitas

Copyright (c) 2024 Hanief Abdur Rahman, Bachtiar Syaiful Bachri, Irena Y. Maureen

✉ Corresponding author: Hanief Abdur Rahman

Email Address: abdurrahman.hanief@gmail.com (Jl. Lidah Wetan, Surabaya, Indonesia)

Received 24 June 2024, Accepted 09 July 2024, Published 12 July 2024

DoI: <https://doi.org/10.31004/cendekia.v8i2.3343>

PENDAHULUAN

Berpikir kritis yaitu sebuah keterampilan yang wajib dimiliki setiap individu pada abad ini. Melimpahnya informasi yang tersedia saat ini membuat keterampilan berpikir kritis berperan penting mengolah dan mengevaluasi informasi tersebut. Di bidang pendidikan, keterampilan berpikir kritis membantu peserta didik didalam memecahkan masalah, membuat serta menilai suatu argumen, dan mengambil keputusan, yang pada akhirnya meningkatkan kualitas pemikirannya (Xiaotong, 2022). Dalam konteks pembelajaran statistika, keterampilan berpikir kritis membantu meningkatkan

kemampuan peserta didik dalam menginterpretasikan dan mengevaluasi informasi statistik (MacGillivray, 2023) sehingga mereka bisa memberikan argumen dan evaluasi berbasis data untuk melakukan pengambilan keputusan yang tepat dalam berbagai situasi (Koga, 2022).

Proses mental, teknik, serta representasi yang digunakan manusia untuk mengambil informasi baru, mengatasi masalah, sampai pada kesimpulan (Sternberg, 1986), dan meningkatkan kemungkinan hasil yang diinginkan secara kolektif (Halpern, 1998) disebut sebagai berpikir kritis. Sedangkan menurut Krulik & Rudnick (1995), mempertanyakan, menghubungkan, menguji, dan mengevaluasi seluruh aspek situasi atau masalah, semuanya adalah bagian dari berpikir kritis, yaitu berpikir analitis serta reflektif. Dari pandangan para ahli diatas, bisa disimpulkan berpikir kritis ialah suatu keterampilan yang melibatkan mengajukan pertanyaan-pertanyaan menyelidik, melakukan analisis menyeluruh, memilih data yang relevan, dan merumuskan pemikiran secara cermat memdapat pemahaman yang tepat dan komprehensif terhadap suatu konsep.

Meningkatkan kemampuan berpikir kritis telah muncul sebagai tujuan utama dalam kurikulum sekolah secara global (Innabi & El Sheikh, 2007; Mason et al., 2010; NCTM, 2000). Di Indonesia, berpikir kritis ialah suatu dimensi dari profil pelajar pancasila yang harus dikembangkan dalam implementasi kurikulum merdeka. Berdasarkan data dari Badan Standar, Kurikulum, dan Asesmen Pendidikan (BSKAP, 2023) Kementerian Pendidikan, Kebudayaan, Riset, serta teknologi, pada saat ini sudah lebih dari 70 persen satuan pendidikan di seluruh Indonesia telah menerapkan Kurikulum Merdeka Belajar. Era merdeka belajar memberikan tantangan besar bagi guru untuk menyelenggarakan pembelajaran yang melatih peserta didik berpikir kritis.

Karena pentingnya keterampilan tersebut, maka pendidik wajib menyadari kemampuan berpikir kritis siswanya. Dengan memahami sejauh mana kemampuan berpikir kritis siswa, diharapkan pendidik bisa menilai kemampuan berpikir kritis siswa dan membuat rencana pembelajaran yang bisa membantunya berkembang. Diperlukan instrumen tes yang dapat diandalkan dan asli untuk mengukur kompetensi tersebut. Sayangnya, hanya ada sedikit penelitian yang dilakukan untuk menciptakan tes yang andal dan valid untuk mengukur kemampuan berpikir kritis siswa ketika membaca konten statistik. Dari *preliminary research* yang telah dilakukan di SMA Kemala Bhayangkari 3 Porong, ditemukan instrumen tes yang dipakai guru belum dapat mengukur kemampuan berpikir kritis peserta didik. Ini terlihat dari instrumen tes dan kisi-kisi yang telah dibuat guru. Berdasarkan teori taksonomi Bloom, keterampilan berpikir kritis termasuk ke dalam domain C4 hingga C6 (menganalisis, mengevaluasi, dan mencipta), namun instrumen tes yang ada sebagian besar hanya mengukur kemampuan berpikir pada domain C1 hingga C3 (mengingat, memahami, dan menerapkan). Dari latar belakang tersebut, penelitian ini bertujuan merancang dan mengembangkan instrumen tes yang valid serta reliabel untuk mengukur keterampilan berpikir kritis peserta didik pada materi statistika.

METODE

Dengan model 4D Thiagarajan (1974) yang mencakup empat tahap: define, design, develop, dan disseminate, penelitian ini adalah penelitian pengembangan. Pendekatan metodis model 4D mengarah pada pemilihannya. Namun penelitian ini hanya diselesaikan sampai tahap Develop agar tercipta instrumen tes yang valid, reliabel, dan mempunyai tingkat kesukaran yang baik serta mampu membedakan karena keterbatasan peneliti.

Pada tahap pendefinisian ditentukan tujuan pembelajaran, dilakukan analisis front-end, analisis siswa, dan analisis ide. Desain ialah fase berikutnya. Tugas yang diselesaikan pada tahap desain meliputi pembuatan kerangka instrumen tes berdasarkan temuan analisis pada tahap pendefinisian dan mengidentifikasi penanda kemampuan berpikir kritis yang akan dinilai. Lima indikator—interpretasi, analisis, inferensi, penilaian, dan penjelasan—yang dikemukakan Facione (2011), dipakai pada penelitian ini untuk mengukur kemampuan berpikir siswa. Hasil kerangka instrumen tersebut yang akan menjadi pedoman dalam mengembangkan butir soal. Pada tahap berikutnya yaitu develop, dilakukan uji validitas isi terhadap rancangan instrumen tes yang telah dibuat. Validitas isi dinilai para ahli menggunakan angket, kemudian dianalisis dengan menentukan nilai *Content Validity Index* (CVI) dari tiap aspek penilaian (i-CVI) maupun secara keseluruhan (s-CVI) (Lynn, 1986). Setelah dilakukan penilaian validator, 38 siswa kelas XI SMA Kemala Bhayangkari 3 Porong mengikuti percobaan singkat. Tujuan dari uji coba kecil ini ialah untuk mengumpulkan informasi mengenai validitas konstruk, reliabilitas, daya pembeda, serta tingkat kesukaran instrumen tes.

Uji validitas konstruk dilakukan dengan cara mencari nilai koefisien korelasi Pearson. Sedangkan reliabilitas instrumen diuji dengan mencari nilai Cronbach's Alpha (Hamzah, 2014). Kemampuan instrumen tes dalam membedakan siswa berkemampuan tinggi serta rendah diukur dengan menggunakan perhitungan daya diferensial. Dengan membandingkan selisih mean kelompok atas dan mean kelompok bawah dengan skor maksimum setiap item, maka ditentukan kekuatan diferensial (Boopathiraj & Chellamani, 2013). Sementara itu, nilai rata-rata dan nilai maksimal tiap soal dibandingkan mengetahui tingkat kesukaran soal (Johari et al., 2011).

HASIL DAN DISKUSI

Pengembangan instrumen tes ini melalui 3 tahap yakni define, design, dan develop. Berikut ialah hasil dari masing-masing tahapan

Define

Analisis front-end: “Pengembangan instrumen tes untuk mengukur kemampuan berpikir kritis pada materi statistika” didasarkan pada beberapa faktor, antara lain pentingnya menentukan kemampuan berpikir kritis siswa dan minimnya penelitian pada pokok bahasan pembuatan tes. yang bisa mengukur kemampuan berpikir kritis siswa. Soal dan grid yang dikembangkan memperlihatkan penilaian guru dan pembelajaran statistika siswa belum dapat mengukur kemampuan berpikir kritis siswa.

Analisis peserta didik: Subjek penelitian yang menjadi sasaran ialah peserta didik kelas XI SMA Kemala Bhayangkari 3 Porong dengan rentang usia antara 15 – 17 tahun. Dilihat dari teori perkembangan kognitif, teori Piaget menempatkan kelompok usia ini ditahap operasional formal. Dalam tahap ini, peserta didik bisa melakukan penalaran deduktif (Jeon et al., 2020), berpikir abstrak (Wright & Kutcher, 2016), membuat perencanaan secara sistematis untuk menyelesaikan masalah, serta membuat hipotesis (Senan, 2013).

Analisis konsep: Materi yang dipilih ialah Statistika dengan submateri Persamaan Regresi Linear dan Koefisien Korelasi. Jika dibandingkan dengan kurikulum sebelumnya, materi ini adalah materi yang baru muncul di Kurikulum Merdeka. Saat menguji hubungan antara dua variabel atau lebih, koefisien korelasi dan persamaan regresi linier digunakan. Gagasan ini dapat digunakan untuk menentukan apakah faktor-faktor tersebut mempunyai hubungan sebab-akibat atau hanya hubungan korelasional.

Design

Output yang dihasilkan pada tahap ini berupa rancangan butir soal yang bisa mengukur ketercapaian tujuan pembelajaran serta keterampilan berpikir kritis peserta didik.

Tabel 1 Rancangan Instrumen Tes Berpikir Kritis

<p>Tujuan pembelajaran:</p> <ol style="list-style-type: none"> 1) Peserta didik bisa menyelesaikan masalah terkait hubungan antara dua variabel numerikal menggunakan diagram scatter. 2) Peserta didik bisa menyelesaikan masalah tentang prediksi nilai suatu variabel menggunakan konsep persamaan garis regresi linear sederhana. 3) Peserta didik bisa menyelesaikan masalah terkait kekuatan hubungan antara dua data bivariat menggunakan konsep koefisien korelasi. 	
<p>ILUSTRASI (Untuk Soal No. 1 – 3)</p> <p>Grafik di bawah ini memperlihatkan data tentang angka partisipasi sekolah usia 16-18 tahun dan persentase penduduk miskin menurut kabupaten/kota yang ada provinsi Jawa Timur (sumber: jatim.bps.go.id)</p> <div style="text-align: center;"> <p>GRAFIK HUBUNGAN ANGKA PARTISIPASI SEKOLAH USIA 16-18 TAHUN DENGAN PERSENTASE PENDUDUK MISKIN TIAP KABUPATEN/KOTA DI PROVINSI JAWA TIMUR</p> </div>	
<p>1. Berdasarkan grafik diatas, apakah terdapat hubungan antara angka partisipasi sekolah dengan persentase penduduk miskin? Bagaimanakah hubungannya, positif atau negatif? Seberapa kuat hubungan antara kedua variabel tersebut? (Skor maks = 20)</p>	<p>Indikator berpikir kritis: Interpretasi Deskriptor: Memahami dan menyatakan makna dari berbagai macam informasi atau data yang diberikan</p>

<p>2. Apakah grafik diatas dapat dijadikan dasar dalam membuat suatu kebijakan dalam rangka mengurangi persentase penduduk miskin yang ada provinsi jawa timur? Analisislah kebijakan apa yang akan anda buat apabila anda menjadi gubernur untuk mengurangi persentase penduduk miskin? Jelaskan argumentasimu! (Skor maks = 20)</p>	<p>Indikator berpikir kritis: Analisis Deskriptor: Mengidentifikasi hubungan antara pertanyaan, pernyataan, atau konsep serta menjelaskan alasan dan opini.</p>																																							
<p>3. Apabila di lapangan ditemukan data suatu kabupaten memiliki nilai APS yang tinggi akan tetapi persentase penduduk miskin dikabupaten tersebut juga tinggi, buatlah dugaan tentang faktor-faktor lain yang mungkin mempengaruhi tingginya persentase penduduk miskin dikabupaten tersebut! (Skor maks = 20)</p>	<p>Indikator berpikir kritis: Inferensi Deskriptor: Membuat hipotesis dan mempertimbangkan informasi lain yang relevan</p>																																							
<p>ILUSTRASI (Untuk Soal No. 4 – 5) Pak Anton adalah seorang pengusaha UMKM (Usaha Mikro Kecil dan Menengah). Agar usahanya dapat berkembang, ia ingin mengevaluasi efektivitas biaya iklan yang telah dikeluarkan dalam meningkatkan penjualan produk. Data biaya iklan dan volume penjualan produk selama satu tahun terakhir terlihat dalam data berikut.</p>																																								
<table border="1"> <thead> <tr> <th>Bulan</th> <th>Biaya Promosi (Rp)</th> <th>Banyak Barang yang Terjual (pcs)</th> </tr> </thead> <tbody> <tr><td>Januari</td><td>Rp23.600</td><td>130</td></tr> <tr><td>Februari</td><td>Rp20.150</td><td>84</td></tr> <tr><td>Maret</td><td>Rp46.600</td><td>291</td></tr> <tr><td>April</td><td>Rp30.500</td><td>164</td></tr> <tr><td>Mei</td><td>Rp51.200</td><td>210</td></tr> <tr><td>Juni</td><td>Rp55.800</td><td>268</td></tr> <tr><td>Juli</td><td>Rp32.800</td><td>107</td></tr> <tr><td>Agustus</td><td>Rp38.550</td><td>153</td></tr> <tr><td>September</td><td>Rp36.250</td><td>130</td></tr> <tr><td>Oktober</td><td>Rp22.450</td><td>107</td></tr> <tr><td>November</td><td>Rp11.350</td><td>94</td></tr> <tr><td>Desember</td><td>Rp63.500</td><td>252</td></tr> </tbody> </table>		Bulan	Biaya Promosi (Rp)	Banyak Barang yang Terjual (pcs)	Januari	Rp23.600	130	Februari	Rp20.150	84	Maret	Rp46.600	291	April	Rp30.500	164	Mei	Rp51.200	210	Juni	Rp55.800	268	Juli	Rp32.800	107	Agustus	Rp38.550	153	September	Rp36.250	130	Oktober	Rp22.450	107	November	Rp11.350	94	Desember	Rp63.500	252
Bulan	Biaya Promosi (Rp)	Banyak Barang yang Terjual (pcs)																																						
Januari	Rp23.600	130																																						
Februari	Rp20.150	84																																						
Maret	Rp46.600	291																																						
April	Rp30.500	164																																						
Mei	Rp51.200	210																																						
Juni	Rp55.800	268																																						
Juli	Rp32.800	107																																						
Agustus	Rp38.550	153																																						
September	Rp36.250	130																																						
Oktober	Rp22.450	107																																						
November	Rp11.350	94																																						
Desember	Rp63.500	252																																						
<p>4. “Tidak ada hubungan antara biaya promosi dengan volume penjualan”. Analisislah kebenaran pernyataan tersebut! Berikanlah argumentasi ataupun grafik yang mendukung! (Skor maks = 20)</p>	<p>Indikator berpikir kritis: Evaluasi Deskriptor: Menilai kebenaran atau validitas dari suatu pernyataan.</p>																																							
<p>5. Apabila Pak Anton menginginkan penjualan produknya dalam satu bulan meningkat menjadi 500 pcs, jelaskan langkah-langkah yang harus anda lakukan memprediksi biaya optimal yang harus ia keluarkan untuk biaya iklan! (Skor maks = 20)</p>	<p>Indikator berpikir kritis: Eksplanasi Deskriptor: Memberikan kesimpulan atau solusi dari suatu permasalahan disertai langkah-langkahnya.</p>																																							

Develop

Pada tahap ini dilakukan uji validitas dan uji reliabilitas. Satu orang dosen ahli teknologi pendidikan dan satu orang dosen ahli matematika diminta menguji validitas isi instrumen tes yang

telah dirancang. Hasil dari uji validitas isi validator bisa terlihat dalam tabel berikut.

Tabel 2 Data Hasil Uji Validitas Isi

Aspek Penilaian	Validator		i-CVI
	1	2	
Kesesuaian permasalahan yang disajikan dengan tujuan pembelajaran	1	1	1
Permasalahan bisa mengukur keterampilan berpikir kritis peserta didik	0,75	1	0,875
Kejelasan maksud dari permasalahan yang disajikan	1	1	1
Kemungkinan permasalahan dapat terselesaikan	1	1	1
Penggunaan bahasa	0,75	1	0,875
s-CVI	0,95		

Menurut Davis (1992), suatu instrumen dikatakan valid apabila memiliki nilai CVI minimal 0,83. Berdasarkan Tabel 2, didapat hasil nilai CVI dari masing-masing aspek penilaian (i-CVI) berkisar antara 0,875 – 1 dan nilai CVI secara keseluruhan (s-CVI) senilai 0,95. Karena nilai tersebut > 0,83, maka bisa disimpulkan instrumen tes yang dikembangkan sudah memenuhi kriteria validitas isi.

Setelah instrumen tes yang dirancang dikatakan valid validator, instrumen tersebut diujicobakan kepada 38 orang peserta didik kelas XI SMA Kemala Bhayangkari 3 Porong. Hasil ujicoba tersebut kemudian dianalisis untuk menguji validitas konstruk dari setiap butir soal. Hasil uji validitas konstruk dari setiap butir soal terlihat dalam tabel dibawah.

Tabel 3 Hasil Analisis Uji Validitas Konstruk

No Butir Soal	r_{hitung}	r_{tabel}	Kriteria
1	0,620	0,320	Valid
2	0,615	0,320	Valid
3	0,649	0,320	Valid
4	0,671	0,320	Valid
5	0,697	0,320	Valid

Menurut Sugiyono (2018), kriteria valid terpenuhi apabila $r_{hitung} > r_{tabel}$. Berdasarkan tabel nilai r Product Moment dengan taraf signifikansi senilai 5% dan jumlah responden sebanyak 38 orang, didapat nilai $r_{tabel} = 0,320$. Berdasarkan hasil analisis, kelima butir soal mendapat nilai $r_{hitung} > r_{tabel}$. Maka bisa dikatakan semua butir soal yang dikembangkan sudah memenuhi kriteria valid.

Data nilai hasil ujicoba kemudian dianalisis kembali untuk mencari reliabilitas dari butir-butir soal yang sudah dikatakan valid. Dari hasil analisis, didapat nilai Cronbach's Alpha senilai 0,656. Menurut Taber (2018), nilai Cronbach's Alpha antara 0.60 dan 0.70 memperlihatkan tingkat reliabilitas suatu instrumen masih dapat diterima.

Analisis berikutnya yaitu daya beda serta tingkat kesukaran soal. Hasil analisis terlihat dalam Tabel 4.

Tabel 4 Hasil Analisis Daya Beda serta tingkat Kesukaran

No Butir Soal	Indeks Daya Beda	Indeks Tingkat Kesukaran
1	0,36	0,59
2	0,42	0,53
3	0,42	0,48
4	0,36	0,39
5	0,46	0,36

Sebuah butir soal dikatakan baik apabila memiliki indeks daya beda diatas 0,35 (Singh et al., 2013) dan indeks tingkat kesukaran diantara 0,3 – 0,8 (Johari et al., 2011). Soal yang terlalu mudah akan mengakibatkan peserta didik tidak termotivasi, sementara soal yang terlalu susah akan mengakibatkan peserta didik putus asa. Berdasarkan hasil analisis, seluruh butir soal memiliki indeks daya beda $> 0,35$. Ini memperlihatkan instrumen tes yang dikembangkan memiliki kemampuan membedakan antara keterampilan berpikir kritis peserta didik yang tinggi dan yang rendah. Dari Tabel 4 juga terlihat kelima butir soal seluruhnya memiliki indeks tingkat kesukaran yang ada dalam interval antara 0,3 – 0,8. Ini berarti instrumen tes yang dikembangkan memiliki tingkat kesulitan sedang.

KESIMPULAN

Berdasarkan hasil penelitian, bisa disimpulkan 5 butir soal instrumen tes untuk mengukur keterampilan berpikir kritis pada materi statistika yang telah dikembangkan sudah memenuhi kriteria valid, reliabel, memiliki daya beda yang baik, serta tingkat kesukaran sedang. Bagaimanapun, yang menjadi subjek uji coba pada penelitian ini terbatas pada peserta didik kelas XI di SMA Kemala Bhayangkari 3 Porong. karena itu, perlu adanya penelitian lanjutan dengan subjek yang memiliki latar belakang pendidikan berbeda agar instrumen tes yang dikembangkan bisa diimplementasikan guna mengukur keterampilan berpikir kritis peserta didik dengan latar belakang yang lebih luas.

REFERENSI

- Boopathiraj, C., & Chellamani, K. (2013). Analysis of Test Items on Difficulty Level and Discrimination Index in the Test for Research in Education. *International Journal of Social Science & Interdisciplinary Research*, 2(2), 189–193.
- BSKAP. (2023). *Keputusan Kepala Badan Standar, Kurikulum, dan Asesmen Pendidikan Kementerian Pendidikan, Kebudayaan, Riset, serta teknologi Nomor 028/H/KR/2023 tentang Satuan Pendidikan Pelaksana Implementasi Kurikulum Merdeka pada Tahun Ajaran 2023/2024*. Kementerian Pendidikan, Kebudayaan, Riset, serta teknologi.
- Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, 5(4), 194–197. [https://doi.org/10.1016/S0897-1897\(05\)80008-4](https://doi.org/10.1016/S0897-1897(05)80008-4)
- Facione, P. a. (2011). Critical Thinking : What It Is and Why It Counts. In *Insight assessment* (Issue

- ISBN 13: 978-1-891557-07-1.). Insight Assessment. <https://www.insightassessment.com/CT-Resources/Teaching-For-and-About-Critical-Thinking/Critical-Thinking-What-It-Is-and-Why-It-Counts/Critical-Thinking-What-It-Is-and-Why-It-Counts-PDF>
- Halpern, D. F. (1998). Teaching Critical Thinking for Transfer Across Domains: Dispositions, Skills, Structure Training, and Metacognitive Monitoring. *American Psychologist*, 53(4), 449–455. <https://doi.org/10.1037/0003-066X.53.4.449>
- Hamzah, A. (2014). *Evaluasi Pembelajaran Matematika*. Raja Grafindo Persada.
- Innabi, H., & El Sheikh, O. (2007). The change in mathematics teachers' perceptions of critical thinking after 15 years of educational reform in Jordan. *Educational Studies in Mathematics*, 64(1), 45–68. <https://doi.org/10.1007/s10649-005-9017-x>
- Jeon, M., Draney, K., Wilson, M., & Sun, Y. (2020). Investigation of adolescents' developmental stages in deductive reasoning: An application of a specialized confirmatory mixture IRT approach. *Behavior Research Methods*, 52(1), 224–235. <https://doi.org/10.3758/s13428-019-01221-5>
- Johari, J., Sahari, J., Wahab, D. A., Abdullah, S., Abdullah, S., Omar, M. Z., & Muhamad, N. (2011). Difficulty Index of Examinations and Their Relation to the Achievement of Programme Outcomes. *Procedia - Social and Behavioral Sciences*, 18, 71–80. <https://doi.org/https://doi.org/10.1016/j.sbspro.2011.05.011>
- Koga, S. (2022). Characteristics of statistical literacy skills from the perspective of critical thinking. *Teaching Statistics*, 44(2), 59–67. <https://doi.org/10.1111/test.12302>
- Krulik, S., & Rudnick, J. A. (1995). *The New Sourcebook for Teaching Reasoning and Problem Solving in Elementary School. A Longwood Professional Book*. Allyn dan Bacon.
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382–386. <https://doi.org/10.1097/00006199-198611000-00017>
- MacGillivray, H. (2023). C Oswald George Prize Announcement 2022. *Teaching Statistics*, 45(1), 57–57. <https://doi.org/10.1111/test.12327>
- Mason, J., Burton, L., & Stacey, K. (2010). *Thinking Mathematically* (2nd ed.). Pearson Education Limited.
- NCTM. (2000). *Principles and standards for school mathematics (Vol. 1)* (Vol. 1). National Council of Teachers of Mathematics. <http://ezproxy.library.wisc.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED429844&site=ehost-live>
- Senan, D. C. (2013). Enhancing educational effectiveness in physics through cognitive development model: An investigation in formal operational stage. *The International Journal of Engineering and Science*, 2(8), 36–41. <http://theijes.com/papers/v2-i8/Part.3/G0283036041.pdf>
- Singh, D., Singh, T., & Gupta, P. (2013). Test and Item Analysis. In *Principles of Medical Education* (pp. 108–108). https://doi.org/10.5005/jp/books/11904_17

- Sternberg, R. J. (1986). *Critical thinking: Its nature, measurement, and improvement*. National Institute of Education.
- Sugiyono. (2018). *Metode Penelitian Kuantitatif, Kualitatif, serta r&D*. Alfabeta.
- Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Thiagarajan, S., Semmel, D. S., & Semmel, M. I. (1974). *Instructional Development for Training Teachers of Exceptional Children: A Sourcebook*.
- Wright, L., & Kutcher, S. (2016). Adolescent Brain Development. In *Colloquium Series on The Developing Brain* (Vol. 5, Issue 1, pp. 1–104). <https://doi.org/10.4199/c00133ed1v01y201602dbr012>
- Xiaotong, C. (2022). Enhancing Motivation to Think Critically. *Journal of Higher Education Research*, 3(4), 389–389. <https://doi.org/10.32629/jher.v3i4.1019>