

PENGEMBANGAN INSTRUMEN TES KEMAMPUAN BERPIKIR KRITIS PADA MATERI SISTEM PERNAPASAN

Kurniahtunnisa*¹⁾, Zusje Wiesje Merry Warouw²⁾, Musma Rukmana³⁾

^{1,2)}Pendidikan IPA, Fakultas MIPA dan Kebumian, Universitas Negeri Manado

³⁾Pendidikan Biologi, Fakultas MIPA dan Kebumian, Universitas Negeri Manado, Sulawesi Utara, Indonesia

*e-mail: kurniahtunnisa@unima.ac.id*¹⁾, *zusyewarouw@unima.ac.id*²⁾, *musmarukmana@unima.ac.id*³⁾
**Corresponding author*

Received: Dec. 27th, 2023; Revised: Jan. 23th, 2024; Accepted: Feb. 24th, 2024; Published: April 29th, 2024

ABSTRAK

Kemampuan berpikir kritis merupakan aspek esensial dalam pembelajaran IPA, yang menjadi fondasi bagi siswa untuk mengambil keputusan dan memecahkan permasalahan kontekstual. Studi ini bertujuan untuk mengembangkan instrumen tes untuk mengukur kemampuan berpikir kritis siswa pada topik sistem pernapasan. Metode yang diterapkan adalah Penelitian dan Pengembangan (R&D). Hasil validasi oleh dua validator mengindikasikan bahwa instrumen tes sangat valid pada aspek kesesuaian materi, konstruksi, dan penggunaan bahasa pada soal, dengan rata-rata skor sebesar 86,67. Analisis empiris melalui Anates 4.0 menghasilkan 29 soal valid dan 13 tidak valid. Terdapat 11 soal memiliki daya beda baik, 20 cukup, dan 11 jelek. Kesulitan soal bervariasi; 9 mudah, 18 sedang, dan 15 sukar hingga sangat sukar. Uji reliabilitas diperoleh nilai 0,70 yang berarti instrumen reliabel. Berdasarkan validitas isi dan empiris, 25 soal dari 42 soal yang dikembangkan dapat digunakan dan efektif mengukur kemampuan berpikir kritis siswa, dengan hasil rata-rata menunjukkan siswa berada pada kategori cukup kritis.

Kata Kunci: instrumen soal; kemampuan berpikir kritis; validitas empiris; validitas isi

PENDAHULUAN

Dalam dunia pendidikan, diketahui bahwa kemampuan berpikir kritis sangat vital, terutama dalam proses memahami Ilmu Pengetahuan Alam (IPA) terpadu. Mengintegrasikan berbagai aspek IPA memerlukan pemahaman yang mendalam dan melibatkan penalaran yang cermat - dua elemen yang merupakan intisari dari berpikir kritis (Paul & Elder, 2006). Mengembangkan pemahaman akan konsep-konsep dalam IPA tidak cukup dengan hanya memahami faktanya saja, namun juga memahami hubungan antara fakta-fakta tersebut, dan bagaimana mereka saling berkaitan dalam konteks yang lebih besar (National Research Council, 2012).

Inilah yang dimaksud dengan pemahaman mendalam dan penalaran cermat atau berpikir kritis. Selain itu, berpikir kritis tidak hanya membantu merefleksikan dan menganalisis informasi tapi juga berguna dalam melatih siswa menjadi pemikir independent dan kreatif yang dapat memberikan kontribusi signifikan dalam pembelajaran IPA terpadu (Facione & Facione, 2007). Salah satu tujuan utama pendidikan adalah membekali siswa untuk menjadi penuntut ilmu sepanjang hayat - dan berpikir kritis adalah keterampilan esensial dalam proses tersebut.

Kemampuan untuk berpikir secara kritis adalah aspek fundamental yang memungkinkan seseorang untuk mengatasi masalah yang ditemuinya. Berpikir kritis sangat penting karena ini menjadi dasar bagi siswa dalam membuat keputusan dan melakukan tindakan atas tantangan yang dialami dalam kehidupan sehari-hari. Kemampuan berpikir kritis adalah salah satu dari kemampuan kognitif tertinggi yang diakui secara global dan dianggap sebagai kompetensi utama dalam dunia pendidikan, terutama dalam bidang sains dan teknologi Pedrosa-de-Jesus et al. (2014). Memiliki keterampilan berpikir kritis memiliki dampak positif pada kemampuan siswa dalam memecahkan masalah dan meningkatkan kepercayaan diri mereka (Azizi et al., 2018). Perspektif ini didukung oleh pandangan Addy et al. (2014) yang menyatakan adanya hubungan kuat antara kemampuan berpikir kritis dan pencapaian akademik siswa. Sehingga kemampuan berpikir kritis menjadi penting untuk diukur dalam kurikulum pendidikan.

Mengukur kemampuan siswa dalam berpikir kritis didasari oleh dua alasan utama yang signifikan. Pertama, berpikir kritis memainkan peran yang vital dalam mentransformasikan pengetahuan pasif siswa menjadi pemahaman aktif yang diperlukan untuk berinteraksi secara efektif dengan lingkungan sekitarnya. Dengan memiliki kemampuan berpikir kritis, siswa dapat menginterpretasikan konsep-konsep dengan lebih baik, mentransformasikan ide-ide menjadi tulisan, dan menjadi lebih independen dalam proses pembelajaran mereka (Mirra et al., 2018). Berikutnya, kemampuan berpikir kritis adalah keterampilan yang berharga untuk hampir semua aspek kehidupan, termasuk situasi profesional dan dalam menyelesaikan

tantangan pribadi sehari-hari. Menghadapi masa depan, siswa dituntut untuk membuat keputusan dan menyelesaikan masalah dalam berbagai domain kehidupan mereka, dan kemampuan berpikir kritis menjadi alat penting dalam membantu mereka melakukan hal tersebut (Duron et al., 2006). Sehingga, sangat penting untuk menciptakan alat evaluasi yang efektif dan tepat guna untuk mengevaluasi kemampuan berpikir kritis siswa.

Mengukur kemampuan berpikir kritis memang memiliki tantangan tersendiri. Proses ini memerlukan alat evaluasi yang akurat yang mampu merangsang siswa untuk memanfaatkan pengetahuannya dalam menganalisis berbagai situasi dan masalah. Alat evaluasi ini harus dapat memverifikasi bahwa siswa mempunyai pemahaman yang baik terhadap konsep-konsep dasar dan mengukur dengan cermat sejauh mana siswa mampu mengintegrasikan berpikir kritis ke dalam proses pembelajarannya (Huber & Kuncel, 2016). Instrumen yang berkualitas adalah instrumen yang telah memenuhi standar validitas dan reliabilitas. Validitas membantu memastikan hasil yang akurat dan relevan, sementara reliabilitas membantu memastikan hasil yang konsisten dan dapat diandalkan (Heale & Twycross, 2015). Oleh karena itu, dalam mengembangkan atau memilih instrumen penilaian, penting untuk mempertimbangkan baik validitas maupun reliabilitasnya. Sehingga manajemen dan pengembangan instrumen evaluasi yang efektif menjadi sangat penting sebagai upaya untuk membuat proses belajar mengajar menjadi lebih bermakna dan holistik (Duron et al., 2006). Penelitian ini mengkaji pengembangan instrumen soal yang valid dan reliabel untuk mendapatkan

data yang akurat terkait profil kemampuan berpikir kritis siswa.

METODE

Dalam penelitian ini, digunakan pendekatan model ADDIE yang terdiri dari lima langkah, yakni Analisis, Desain, Pengembangan, Implementasi, dan Evaluasi. Tahap analisis mencakup studi pendahuluan terkait kebutuhan alat evaluasi guna mengevaluasi kemampuan berpikir kritis siswa pada materi sistem pernapasan. Prosedur ini melibatkan studi literatur dan survei awal yang melibatkan guru dan siswa. Tahap desain penulis merancang

desain awal instrumen soal berdasarkan hasil analisis dan penelitian teoritis terkait sistem pernapasan dan berpikir kritis. Pada tahap pengembangan penulis menyusun 42 nomor soal berdasarkan kisi-kisi soal serta melakukan pengecekan validitas soal oleh ahli materi dan ahli evaluasi. Soal dikembangkan berdasarkan 16 sub-indikator berpikir kritis dari Ennis (2016) yang ditampilkan pada Tabel 1.

Tabel 1. Indikator dan Sub-Indikator Kemampuan Berpikir Kritis

No.	Indikator	Sub Indikator
1	Memberikan penjelasan sederhana	Memfokuskan pertanyaan Menganalisis argumen Bertanya dan menjawab pertanyaan tentang suatu penjelasan Memahami dan menggunakan grafik dan matematika dasar
2.	Membangun keterampilan dasar	Mempertimbangkan apakah sumber dapat dipercaya atau tidak Mengobservasi dan mempertimbangkan hasil observasi Menggunakan pengetahuan atau prosedur yang sudah ada
3.	Menyimpulkan	Mendeduksi dan mempertimbangkan hasil deduksi Menginduksi dan mempertimbangkan hasil induksi Membuat dan menentukan hasil pertimbangan
4.	Memberikan penjelasan lanjut	Mendefenisikan istilah dan mempertimbangkan suatu defenisi Mengidentifikasi asumsi Sepakat dengan label kesalahan Menangani dalih dengan tepat Berpikir secara proporsional
5.	Memberikan alternatif pemecahan masalah	Menggunakan strategi retorikal

Pada tahap Implementasi, instrumen diujicoba skala terbatas terhadap 36 siswa di SMA Negeri 12 Semarang, guna mengevaluasi efektivitas soal yang telah dirancang. Soal dianalisis berdasarkan tingkat kesulitan, pengecoh/distraktor, validitas, dan reliabilitas menggunakan

aplikasi Anates. Tahap Evaluasi melibatkan interpretasi data uji coba dan penilaian kembali desain dan pengembangan soal, serta melakukan perbaikan soal jika diperlukan. Interpretasi skor validitas isi menggunakan kriteria yang ditampilkan pada Tabel 2.

Tabel 2. Kriteria Validitas Isi Instrumen Soal

Presentase Skor (%)	Kriteria
$81,25 < P \leq 100$	Sangat valid
$62,50 < P \leq 81,25$	Valid
$43,75 < P \leq 62,50$	Tidak valid
$25,00 < P \leq 43,75$	Sangat tidak valid

Sebanyak 25 soal yang telah valid diujicobakan skala luas kepada 106 siswa di SMA Negeri 16 Semarang untuk mengukur kemampuan berpikir kritis siswa.

HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan instrumen soal untuk mengukur kemampuan berpikir kritis siswa. Instrumen ini telah divalidasi oleh ahli materi dari Guru Biologi SMA Negeri 16 Semarang dan ahli evaluasi dari Dosen Universitas Negeri Semarang. Instrumen tersebut selanjutnya telah diuji dan dianalisis menggunakan aplikasi Anates. Rekapitulasi hasil validitas isi instrumen soal ditampilkan pada Tabel 4.

Tabel 4. Rekapitulasi Validitas Isi Instrumen Soal

Validator	Persentase	Kriteria
Ahli Materi	88,33	Sangat Valid
Ahli Evaluasi	85	Sangat Valid
Rata-rata	86,67	Sangat Valid

Berdasarkan Tabel 4, rata-rata instrumen soal menunjukkan tingkat yang sangat valid. Kevalidan ini tercapai karena penilaian oleh validator pada aspek materi, konstruksi, dan bahasa dianggap sangat sesuai dan efektif dalam mengukur kemampuan berpikir kritis. Hasil penelitian yang dipublikasikan oleh Salsabila &

Pencapaian siswa diinterpretasi menjadi skor kemampuan berpikir kritis dengan menggunakan kriteria yang ditampilkan pada Tabel 3.

Tabel 3. Kriteria kemampuan berpikir kritis

Skor	Kriteria
$81,26\% < \text{nilai} \leq 100\%$	Sangat kritis
$62,51\% < \text{nilai} \leq 81,25\%$	Kritis
$43,76\% < \text{nilai} \leq 62,50\%$	Cukup kritis
$25,01\% < \text{nilai} \leq 43,75\%$	Kurang kritis
$00,00\% < \text{nilai} \leq 25\%$	Sangat kurang kritis

Trimulyono, (2023) mengenai pengembangan instrumen soal juga merekomendasikan perlunya validasi dengan melibatkan para ahli untuk meningkatkan validitas dan reliabilitas instrumen. Selain memberikan penilaian, validator juga memberikan komentar dan saran untuk meningkatkan kualitas instrumen soal. Tabel 5 menampilkan komentar dan saran dari validator.

Tabel 5. Komentar dan Saran Hasil Validasi Isi

Butir Soal	Komentar dan Saran
15	Pilihan jawaban pada soal belum ada pilihan yang tepat
Cek seluruh butir soal	Membuat soal evaluasi dengan jumlah yang sesuai presentase pada soal C3 sampai C6 dan indikator berpikir kritis serta tambahkan alasan untuk mengantisipasi asal jawab oleh siswa sehingga diubah menjadi pilihan ganda beralasan.
4, 25, 26	Butir soal disesuaikan dengan indikator butir soal
21, 27	Tambahkan sumber artikel yang dicantumkan pada soal

Instrumen soal telah diperbaiki sesuai masukan dan arahan dari semua validator. Persentase soal telah disesuaikan untuk tingkat mudah, sedang, dan sukar serta telah disesuaikan dengan kriteria berpikir kritis menurut Ennis. Soal telah diubah

menjadi soal pilihan ganda beralasan dan telah dibuatkan rubrik untuk menilai pilihan jawaban dan alasan siswa. Contoh kisi-kisi soal dan contoh soal untuk mengevaluasi kemampuan berpikir kritis siswa yang telah disusun ditampilkan pada Gambar 1 dan 2.

No.	Indikator Pembelajaran	Aspek kemampuan berpikir kritis	Sub indikator kemampuan berpikir kritis	Indikator soal	Nomor Soal	Tingkatan Kognitif Soal				Kunci Jawaban
						C3	C4	C5	C6	
6	Menjelaskan hubungan kondisi udara lingkungan yang tidak bersih terhadap sistem pernapasan	Memberikan penjelasan sederhana	Menganalisis argumen	Diberikan bacaan tentang pendapat dua ilmuwan yang bertentangan, siswa memberikan alasan yang mendukung atau meragukan ilmuwan tersebut	28		√			A
		Membangun keterampilan dasar	Mempertimbangkan apakah sumber dapat dipercaya atau tidak	Diberikan wacana pendapat seorang ahli terkait dengan pencemaran udara, siswa menganalisis argumen tersebut dengan memberikan alasan yang tepat.	27			√		B
		Menyimpulkan	Membuat dan menentukan hasil pertimbangan	Diberikan bacaan tentang pendapat dua ilmuwan yang bertentangan, siswa memberikan pendapat yang membuat hasil penelitian tersebut kurang tepat	30			√		C

Gambar 1. Contoh kisi-kisi instrumen soal yang dikembangkan

RESIKO KESEHATAN?

Bayangkan bahwa kamu tinggal di dekat pabrik kimia besar yang menghasilkan pupuk bagi pertanian. Pada tahun-tahun belakangan ini, terjadi beberapa kasus pada masyarakat di sekitar pabrik. Mereka menderita penyakit pernapasan yang menahun. Masyarakat sekitar percaya bahwa gejala ini disebabkan oleh emisi uap beracun dari pabrik pupuk kimia yang ada di sana. Sebuah pertemuan masyarakat diadakan untuk mendiskusikan potensi bahaya dari pabrik kimia terhadap kesehatan penduduk sekitar. Pada pertemuan ini, para ilmuwan membuat pernyataan berikut.

Pernyataan ilmuwan yang bekerja untuk perusahaan kimia
"Kami telah meneliti kandungan racun dari tanah di daerah sekitar. Kami tidak menemukan adanya bahan kimia beracun di dalam contoh tanah yang kami ambil."

Pernyataan ilmuwan yang bekerja untuk komunitas warga setempat
"Kami telah meneliti jumlah kasus penyakit pernapasan menahun pada warga di sekitar lokasi dan membandingkannya dengan jumlah kasus pada warga yang tinggal jauh dari pabrik kimia. Terdapat lebih banyak kejadian penyakit ini di lokasi dekat dengan pabrik kimia."

28. Pemilik pabrik kimia menggunakan pernyataan ilmuwan yang bekerja untuk perusahaan untuk berdalih bahwa "emisi uap dari pabrik tidak beresiko kesehatan bagi warga sekitar".

Berikut ini alasan yang tepat untuk **meragukan** bahwa pernyataan ilmuwan yang bekerja untuk perusahaan mendukung pendapat si pemilik adalah...

A. sampel yang digunakan pada penelitian bisa jadi belum mewakili seluruh lokasi
 B. memang emisi uap dari pabrik tidak beresiko kesehatan bagi warga sekitar
 C. sampel yang digunakan sudah mewakili seluruh lokasi tetapi tidak ada hubungan emisi uap dengan penyakit pernapasan
 D. penelitian kandungan racun dari tanah sudah mewakili sampel kandungan racun di udara
 E. hasil penelitian di ubah atas permintaan pemilik perusahaan

Gambar 2. Contoh instrumen soal yang dikembangkan

Setelah dilakukan validasi isi, dilanjutkan dengan validasi empiris pada tes kemampuan berpikir kritis. Rekapitulasi

hasil validasi empiris jawaban uji coba skala kecil disajikan pada Tabel 6.

Table 6. Hasil Analisis Jawaban Ujicoba Butir Soal

Analisis Jawaban	Kriteria	Jumlah Soal	Nomor Soal
Tingkat kesukaran	Sangat sukar	2	33,36
	Sukar	13	2,3,5,8,10,14,21,24,29,30,31,34,38
	Sedang	18	1,9,12,13,15,16,17,18,19,20,22,26,28, 35,37,39, 40,41
	Mudah	3	4,25,42
	Sangat mudah	6	6,7,11,23,27,32
Daya beda	Baik	11	1, 9, 10, 14, 16, 20, 22, 26, 29, 30, 37
	Cukup	20	2, 3, 5, 6, 8, 11, 12, 13, 15, 17, 18, 21, 28, 31, 33, 35, 38, 39, 40, 41
	Jelek	11	4, 7, 19, 23, 24, 25, 27, 32, 34, 36, 42
Validitas	Sangat valid	16	1, 2, 3, 8, 9, 10, 11, 14, 16, 18, 21, 22, 29, 30, 31, 41
	Valid	13	5, 6, 12, 13, 20, 26, 28, 34, 35, 36, 37, 38, 40
	Tidak valid	13	4, 7, 15, 17, 19, 23, 24, 25, 27, 32, 33, 39, 42
Uji reliabilitas butir soal	Reliabilitas soal tes dengan rumus belah dua ganjil-genap didapat r_{11} sebesar 0,70 dengan $n = 36$ dan $\alpha = 5\%$.		
Soal yang dapat digunakan untuk ujicoba skala besar	1, 2, 5, 6, 8, 9, 10, 11, 12, 13, 15, 18, 20, 21, 22, 26, 28, 29, 34, 35, 36, 37, 38, 40, dan 41		

Berdasarkan hasil analisis pada Tabel 6 diperoleh 29 soal dengan kriteria valid dan sangat valid serta 13 soal dengan kriteria tidak valid; soal dengan daya beda yang baik 11, cukup 20, dan jelek 11 soal; tingkat kesulitan dengan kriteria mudah 9, sedang 18, sukar dan sangat sukar 15 soal. Evaluasi yang dikembangkan sangat valid. Hal ini terlihat dari penilaian para validator dalam aspek materi, konstruksi, dan bahasa yang sangat baik, sehingga efektif untuk mengukur kemampuan kritis berpikir siswa. Proses pengujian instrumen evaluasi melibatkan validasi isi oleh para pakar, dilanjutkan dengan pengujian validitas empiris pada soal kemampuan berpikir kritis. Proses analisis, yang menggunakan program Anates 4.0, menghasilkan berbagai item tes dengan kualitas yang

berbeda; ada yang valid dan tidak valid, memiliki daya beda yang baik, cukup, maupun jelek, serta tingkat kesulitan berkisar dari sangat mudah, mudah, sedang, hingga sukar.

Dari hasil uji diperoleh 25 soal yang memenuhi syarat digunakan untuk uji coba skala besar. Soal-soal yang ambil untuk uji coba skala besar adalah soal yang sangat valid dan valid, memiliki pengecoh/distraktor yang baik atau cukup baik, dan memiliki tingkat kesulitan mudah, sedang, dan sukar, dan reliabel. Soal dinyatakan reliabel dengan demikian soal tersebut memiliki derajat konsistensi data sehingga mampu mengukur kemampuan berpikir kritis siswa. Hal ini sesuai dengan persyaratan tes yang baik. Kriteria untuk tes

yang baik mencakup validitas, reliabilitas, kemampuan untuk membedakan, dan tingkat kesulitan soal yang sesuai. Kriteria yang paling penting dari semua ini adalah validitas tes. Evaluasi soal yang valid dan reliabel adalah kunci dalam pengukuran kemampuan berpikir kritis siswa. Validitas mensyaratkan bahwa soal mampu mengukur secara tepat apa yang hendak diukur, sementara reliabilitas menegaskan bahwa hasil pengukuran bersifat konsisten dan dapat diandalkan (Arikunto, 2013; Heale & Twycross, 2015).

Berdasarkan hasil analisis soal evaluasi yang dikembangkan valid digunakan untuk mengevaluasi sejauh mana pemikiran kritis siswa. Bagian penting dari validitas ini adalah adanya validasi terhadap soal-soal sebelum diuji coba, dan konfirmasi dari validator mengenai tingginya validitas soal tersebut. Ihsan (2015) menegaskan pentingnya proses validasi isi, yang bertujuan untuk meminimalisir peluang kesalahan dalam penyusunan instrumen dan meningkatkan peluang mendapatkan indeks validitas konstruksi untuk penelitian berikutnya. Sebaliknya, ketika soal tidak memenuhi kriteria yang ditetapkan - dalam hal ini, soal tersebut tidak valid, tidak reliabel, memiliki daya pembeda yang rendah dan tingkat kesulitan yang terlalu ekstrem (terlalu mudah atau terlalu sulit) - perbaikan atau penggantian soal sangat dianjurkan untuk mencapai tujuan pengukuran kompetensi yang sama. Nuswawati et al. (2010) mendukung logika ini, dengan mengemukakan bahwa validitas dan reliabilitas soal sangat mempengaruhi apakah pengukuran pencapaian kompetensi siswa akan akurat atau tidak.

Setelah berhasil memvalidasi instrumen soal, tahap berikutnya adalah melakukan uji coba skala besar. Uji coba tersebut dilakukan pada 106 siswa dari SMAN 16 Semarang, dan hasilnya digunakan untuk mendapatkan data tentang profil kemampuan berpikir kritis siswa. Detail dari hasil tersebut dapat dilihat di Tabel 7.

Tabel 7. Rekap Kemampuan Berpikir Kritis Siswa

Aspek	Nilai
Jumlah Responden	109
Nilai terendah	27
Nilai Tertinggi	60
Standar Deviasi	6,93
Nilai Maksimum	100
Rata-Rata Nilai	56,10 (Cukup Kritis)

Pada Tabel 7 menunjukkan bahwa instrumen soal yang telah dikembangkan efektif mengukur kemampuan berpikir kritis siswa, meskipun mayoritas siswa memperoleh kriteria cukup kritis. Ada beberapa faktor yang mempengaruhi kapasitas berpikir kritis siswa, diantaranya adalah proses pendidikan yang berlangsung di sekolah. Hal ini sejalan dengan gagasan Santos (2017) yang mengatakan bahwa pengembangan kemampuan berpikir kritis siswa di bidang sains kerap terkendala oleh fokus pembelajaran pada penguasaan konten materi dan daya hafal, ketimbang menggali lebih dalam mengenai kreativitas, pemahaman substansial, dan pemikiran kritis siswa.

Dalam konteks ini, kebiasaan siswa dalam mengikuti metode pembelajaran aktif memegang peranan penting untuk meningkatkan potensi pemikiran kritis siswa (Nuryanti et al., 2018). Selain itu, kapasitas siswa untuk menyelesaikan tugas-tugas yang lebih kompleks juga turut

mempengaruhi kemampuan berpikir kritis mereka, sebagian besar siswa belum terlatih dalam menyelesaikan soal-soal pada level HOTS (Asviangga, et al., 2018). Untuk menciptakan lingkungan belajar yang lebih

kondusif, diperlukan perubahan dalam metode pengajaran yang menekankan pada penumbuhkembangan berpikir kritis dan pemahaman konseptual yang lebih dalam (Huber & Kuncel, 2016).

KESIMPULAN

Hasil validasi oleh dua validator mengindikasikan bahwa instrumen tes memperoleh kriteria yang sangat valid pada aspek kesesuaian materi, konstruksi, dan penggunaan bahasa pada soal. Analisis item dilakukan secara empiris menggunakan program Anates, menghasilkan soal-soal dengan validitas tinggi, daya beda yang baik, tingkat kesulitan bervariasi. Uji reliabilitas diperoleh nilai 0,70 yang berarti instrumen reliabel. Berdasarkan validitas isi dan empiris yang berhasil diverifikasi, 25 item soal dari 42 soal yang dikembangkan dinyatakan memenuhi syarat untuk mengukur kemampuan berpikir kritis siswa. Instrumen yang dikembangkan efektif dalam menilai kemampuan berpikir kritis siswa, dengan nilai rata-rata siswa masuk dalam kategori cukup kritis.

UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Pimpinan dan Guru Biologi di SMA Negeri 12 Semarang dan SMA Negeri 16 Semarang yang telah memberi ruang kepada penulis untuk melakukan penelitian.

REFERENSI

Addy, T. M., LePrevost, C., & Stevenson, M. (2014). Thinking Critically in Undergraduate Biology: Flipping the Classroom and Problem-Based Learning. *Double Helix: A Journal of Critical Thinking and Writing*, 2(1), 1–9. <https://doi.org/10.37514/dbh-j.2014>

4.2.1.09

Arikunto, S. (2013). *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara.

Asviangga, A. B., Sunardi, & Trapsilasiwi, D. (2018). Analisis Kemampuan 4C's Siswa dalam Menyelesaikan Soal Matematika Berpikir Tingkat Tinggi. *Kadikma*, 9(1), 17–23.

Azizi, M., Sedaghat, Z., & Direkvand-Moghadam, A. (2018). Effect of Critical Thinking Education on Problem Solving Skills and Self-Esteem in Iranian Female Students. *Journal of Clinical and Diagnostic Research*, 12(1), VC04-VC07. <https://doi.org/10.7860/JCDR/2018/27993.11128>

Duron, R., Limbach, B., & Waugh, W. (2006). Critical Thinking Framework For Any Discipline. *International Journal of Teaching and Learning in Higher Education*, 17(2), 160–166.

Ennis, R. H. (2016). Critical Thinking Across the Curriculum: A Vision. *Topoi*, 37(1), 165–184. <https://doi.org/10.1007/s11245-016-9401-4>

Facione, P. A., & Facione, N. C. (2007). Talking Critical Thinking. *Change: The Magazine of Higher Learning*, 39(2), 38–45. <https://doi.org/10.3200/chng.39.2.38-45>

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-Based Nursing*, 18(3), 66–67. <https://doi.org/10.1136/eb-2015-102129>

Huber, C. R., & Kuncel, N. R. (2016). Does College Teach Critical Thinking? A

- Meta-Analysis. *Review of Educational Research*, 86(2), 431–468. <https://doi.org/10.3102/0034654315605917>
- Ihsan, H. (2015). Validitas Isi Alat Ukur Penelitian: Konsep dan Panduan Penilaiannya. *PEDAGOGIA: Jurnal Ilmu Pendidikan*, 13(3), 266–273. <https://doi.org/10.17509/pedagogia.v13i3.6004>
- Mirra, N., Morrell, E., & Filipiak, D. (2018). From Digital Consumption to Digital Invention: Toward a New Critical Theory and Practice of Multiliteracies. *Theory into Practice*, 57(1), 12–19. <https://doi.org/10.1080/00405841.2017.1390336>
- National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. The National Academies Press.
- Nuryanti, L., Zubaidah, S., & Diantoro, M. (2018). Analisis Kemampuan Berpikir Kritis Siswa SMP. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, 3(2), 155–158.
- Nuswawati, M., Binadja, A., Soeprodjo, & Ifada, K. E. N. (2010). Pengaruh Validitas dan Reliabilitas Butir Soal Ulangan Akhir Semester Bidang Studi Kimia terhadap Pencapaian Kompetensi. *Jurnal Inovasi Pendidikan Kimia*, 4(1), 566–573.
- Paul, R., & Elder, L. (2006). Critical Thinking: The Nature of Critical and Creative Thought, Part II. *Journal of Developmental Education*, 30(3), 36–37. <https://georgefox.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=25122573&scope=site>
- Pedrosa-de-Jesus, H., Moreira, A., Lopes, B., & Watts, M. (2014). So much more than just a list: Exploring the nature of critical questioning in undergraduate sciences. *Research in Science and Technological Education*, 32(2), 115–134. <https://doi.org/10.1080/02635143.2014.902811>
- Salsabila, P. A., & Trimulyono, G. (2023). Pengembangan instrumen HOTS materi virus untuk mengukur kemampuan berpikir kritis siswa kelas X SMA. *BioEdu*, 12(2), 287–297. <https://ejournal.unesa.ac.id/index.php/bioedu>
- Santos, L. F. (2017). The Role of Critical Thinking in Science Education. *Journal of Education and Practice*, 8(20), 159–173.