



Pengembangan Instrumen Tes Keterampilan Berpikir Kritis Peserta Didik pada Materi Pokok Ekologi dan Keanekaragaman Hayati Indonesia

Riska Dina Widian¹, Salma Samputri², Rifda Nur Hikmahwati Arif³, Firs⁴, Fauziatul Ihf⁵, Nadila Pasomba⁶

^{1,2,3,4,5,6}Jurusan Pendidikan IPA, Universitas Negeri Makassar, Makassar, Indonesia

Received: 27 June 2026

Revised: 20 July 2026

Accepted: 30 August 2026

Corresponding Author:

Salma Samputri

salmasamputri@unm.ac.id

© 2026 Kappa Journal is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License



DOI:

<https://doi.org/10.29408/kpj.v10i2.36377>

Abstract: Beragam model, pendekatan, dan strategi pembelajaran yang diterapkan oleh pendidik diterima dengan baik oleh peserta didik. Namun, dalam menilai kemampuan peserta didik, pendidik masih menggunakan tes biasa, belum pernah menggunakan tes berpikir kritis. Pendidik merasakan perlunya pengembangan instrumen tes yang diharapkan dapat lebih efektif dalam mengukur kemampuan peserta didik. Penelitian ini bertujuan untuk mengetahui kualitas perbutir soal ditinjau dari kevalidan, tingkat kesukaran dan daya pembeda soal. Penelitian ini menggunakan metode *Research and Development* (R&D) dengan model ADDIE (*Analysis, Design, Development, Impelementation, dan Evaluation*). Hasil penelitian disimpulkan Hasil validasi isi oleh ahli menunjukkan skor *content validity* sempurna sebesar 1 (kategori valid), sementara uji coba empiris terhadap 10 butir soal uraian menghasilkan 9 soal valid dan layak dipakai (nomor 1, 2, 4, 5, 6, 7, 8, 9, dan 10) karena memiliki tingkat kesukaran yang seimbang (sedang hingga mudah) serta daya pembeda yang memadai (cukup hingga baik). Adapun 1 butir soal dinyatakan tidak valid (nomor 3) akibat nilai korelasi dan daya pembeda yang buruk, sehingga memerlukan revisi total atau penggantian sebelum dapat dimanfaatkan secara optimal.

Keywords: ADDIE; Validitas; Tingkat Kesukaran; Daya Pembeda; Kepraktisan

Pendahuluan

Merdeka belajar saat ini menjadi salah satu program yang telah dirancang oleh pemerintah dan menjadi batu loncatan dalam bidang kualitas pendidikan agar menjadikan para peserta didik baik di tingkat sekolah dasar, menengah pertama, menengah atas, hingga universitas menjadi lulusan dengan kualitas terbaik dalam menghadapi segala tantangan yang lebih kompleks di masa yang akan datang. Merdeka belajar berpusat pada kebebasan bagi tenaga pendidik dan peserta didik. Merdeka belajar

memotivasi dalam membentuk kepribadian semangat merdeka, antara peserta didik dan tenaga pendidik dapat secara merdeka, bebas dan bersinergi dalam menggali ilmu pengetahuan dan ketrampilan serta membentuk karakter diri dalam bersosialisasi dilingkungannya berada (Darl^{is et al.}, 2022).

Pendidikan merupakan salah satu pilar utama dalam pembangunan suatu negara. Untuk mencapai tujuan pendidikan yang optimal, pengembangan kurikulum yang baik menjadi sangat penting. Di Indonesia, Kurikulum Merdeka telah diperkenalkan

How to Cite:

Widian, R.D., Samputri, S., Arif, R.N.H., Firs, F., Ihf, F., & Pasomba, N. (2026). Pengembangan Instrumen Tes Keterampilan Berpikir Kritis Peserta Didik Kelas VII SMP pada Materi Pokok Ekologi dan Keanekaragaman Hayati Indonesia. *Jurnal Kappa*, 10(2), 143-150. <https://doi.org/10.29408/kpj.v10i2.36377>

sebagai upaya untuk meningkatkan kualitas pembelajaran di sekolah. Kebijakan ini mengusung pendekatan yang berbeda dalam penyusunan kurikulum dengan fokus pada pemberdayaan peserta didik dan pengembangan keterampilan abad ke-21 (Tuerah & Jeanne, 2023).

Pembelajaran abad 21 juga mengarahkan peserta didik dalam sebuah pengembangan kemampuan berpikir tingkat tinggi. Kemampuan berpikir tingkat tinggi terbagi lima kemampuan, yaitu: kemampuan berpikir logis, kemampuan berpikir kritis, kemampuan berpikir reflektif, kemampuan berpikir metakognisi, dan kemampuan berpikir reflektif. Proses pembelajaran dijenjang sekolah menengah pertama (SMP) memusatkan peserta didik dalam pengembangan kemampuan berpikir tingkat tinggi (Kurniasari & Hanin, 2022).

Berpikir kritis merupakan keterampilan yang sangat penting dalam konteks keterampilan abad ke-21. Kemampuan berpikir menjadi hal penting dalam mencapai keberhasilan pendidikan khususnya kemampuan berpikir tingkat tinggi atau Higher Order Thinking Skill (HOTS) yang dibutuhkan pada era pendidikan abad 21 ini. Kemampuan berpikir tingkat tinggi mencakup kemampuan berpikir kritis, logis, reflektif (Maulina, 2026).

metakognitif dan berpikir kreatif.

Tingkat kemampuan berpikir kritis dapat diukur dan diketahui melalui asesmen dengan menggunakan instrumen tes berpikir kritis. Instrumen tes berpikir kritis merupakan salah satu solusi yang tepat untuk mengetahui tingkat kemampuan berpikir kritis siswa, karena dengan mengerjakan soal-soal di dalamnya siswa dapat terpacu untuk berpikir secara logis dan juga kritis (Marfhadella, 2021).

Keterampilan berpikir kritis (*Critical Thinking Skills*) merupakan keterampilan fundamental dalam memecahkan masalah. Keterampilan ini penting dimiliki oleh siswa dalam menemukan sumber masalah dan bagaimana mencari dan menemukan solusi yang tepat atas masalah yang dihadapi. Keterampilan berpikir kritis dapat ditanamkan dalam berbagai disiplin ilmu. Guru memegang peranan penting dalam merancang dan mengembangkan program pembelajaran yang lebih terfokus pada pemberdayaan keterampilan ini. Keterampilan berpikir kritis penting bagi siswa agar mampu menilai dan menyimpulkan suatu permasalahan dalam proses pembelajaran (Putri *et al.*, 2023).

Keterampilan berpikir kritis yang dibutuhkan pada abad 21 adalah keterampilan berpikir kritis. Tujuan dari keterampilan berpikir adalah untuk meningkatkan kualitas berpikir. Keterampilan berpikir kritis harus dikembangkan dalam proses pembelajaran.

Keterampilan berpikir kritis menjadi salah satu faktor yang mendukung keberhasilan belajar. Berpikir kritis di pandang sebagai suatu keyakinan yang kuat dan hati-hati dengan maksud untuk mengontraskan system berpikir seseorang yang tidak efektif (Rifda, 2023).

Faktanya kemampuan berpikir kritis di Indonesia belum dapat dikategorikan baik dibandingkan dengan negara lain. Hal ini terbukti berdasarkan hasil uji *Program for International Student Assessment (PISA)* tahun 2022 yang menyampaikan bahwa Indonesia terletak di urutan 69 dari 81 negara peserta. Hasil tes PISA dapat dimanfaatkan sebagai kriteria hitungan tingkatan kemampuan berpikir kritis peserta didik (Aziz & Adi, 2024).

Kemampuan berpikir kritis peserta didik pada pembelajaran IPA memiliki persentase 34,26 % dengan kategori rendah. Oleh karena itu perlunya peningkatan kemampuan berpikir kritis siswa pada pembelajaran IPA. Pembelajaran IPA perlu didukung dengan kemampuan berpikir kritis, karena dalam pembelajaran IPA harus dapat menghubungkan dan memahami konsep yang bersifat mikroskopis dan abstrak, hal ini membutuhkan kemampuan dalam menginterpretasi, menganalisis, dan mengevaluasi (Muzahrain *et al.*, 2024).

Instrumen tes penting dalam mengukur kemampuan dan perkembangan peserta didik. Umumnya alat mengacu pada suatu benda atau perangkat yang memudahkan penyelesaian tugas atau pencapaian tujuan dengan lebih efektif dan efisien. Dalam ranah evaluasi, instrumen bertujuan untuk memperoleh hasil yang lebih akurat serta selaras dengan kenyataan yang dinilai. Dalam bidang penelitian pendidikan, penting untuk mengembangkan instrumen evaluasi yang efektif dan berkualitas tinggi. Hal ini akan memastikan bahwa pengukuran diperoleh bersifat obyektif dan akurat (Kurniawati & Yuyu, 2024).

Dalam menyusun soal diperlukan tahapan-tahapan untuk menganalisis butir soal yang bertujuan untuk mengetahui kualitas soal yang sudah dikembangkan. Validitas, reliabilitas, indeks kesukaran dan daya pembeda merupakan suatu formula yang sangat penting dalam memilih butir soal. Aliditas merupakan bukti dari hasil tes yang akan ditafsirkan sesuai dengan tujuan tes. Fungsi validitas yaitu untuk dijadikan sebagai alat ukur terhadap apa yang ingin di ukur. Daya pembeda merupakan kemampuan butir soal untuk membedakan siswa yang berkemampuan sangat tinggi dengan siswa yang mempunyai kemampuan sangat rendah. Indeks kesukaran merupakan derajat kesukaran suatu butir soal dan dinyatakan dalam bentuk bilangan. Jika soal dikatakan baik, maka soal tersebut tidak tergolong mudah atau tidak terlalu sukar (Sudrajat, 2024).

Berdasarkan hasil survei yang dilakukan oleh Programme for Internasional Student Assesment (PISA) yang mengukur kinerja matematika peserta didik, Indonesia memperoleh skor rata-rata 386 berada di peringkat 63 dari 70 negara yang mengikuti. Hal ini menunjukkan bahwa kemampuan peserta didik di Indonesia dalam berpikir tingkat tinggi masih tergolong rendah, dan harus terus ditingkatkan ke level yang lebih tinggi. Karakteristik soal PISA tersebut menuntut kemampuan peserta didik dalam menganalisis, menalar, dan mengharuskan peserta didik untuk memahami terlebih dahulu maksud soal sebelum menyelesaikan soal yang membutuhkan tahap analisis. Hal ini menunjukkan bahwa kemampuan analisis peserta didik di Indonesia masih tergolong sangat rendah (Mudakir, Joko & Nurma, 2020).

Berdasarkan hasil observasi dan wawancara dengan pendidik di bidang Ilmu Pengetahuan Alam (IPA) di SMPN 27 Makassar mengungkapkan bahwa beragam model, pendekatan, dan strategi pembelajaran yang diterapkan oleh pendidik diterima dengan baik oleh peserta didik. Namun, dalam menilai kemampuan peserta didik, pendidik masih menggunakan tes biasa, belum pernah menggunakan tes berpikir kritis. Pendidik merasakan perlunya pengembangan instrumen tes yang diharapkan dapat lebih efektif dalam mengukur kemampuan peserta didik. Berdasarkan hal tersebut, peneliti menyadari bahwa memiliki instrumen yang valid, tingkat kesukaran, daya pembeda dan kepraktisan yang baik sangat penting untuk menilai kemampuan peserta didik. Oleh karena itu, diperlukan pengembangan dan validasi instrumen tes yang berfokus pada keterampilan berpikir kritis peserta didik.

Penelitian ini menggunakan instrumen tes untuk mengukur serta mengidentifikasi kekuatan dan kelemahan berpikir kritis peserta didik. Instrumen ini diharapkan dapat memberikan gambaran yang tepat mengenai sejauh mana peserta didik mampu menerapkan kemampuan analisis, evaluasi, dan kreativitas dalam menyelesaikan permasalahan sehingga pendidik dapat memberikan intervensi pembelajaran yang lebih tepat sasaran untuk meningkatkan berpikir kritis peserta didik.

Metode

Penelitian ini menggunakan metode *Research and Development* (R&D) dengan model ADDIE (*Analysis, Design, Development, Impelementation, dan Evaluation*). Namun penelitian ini dibatasi hingga tahap *Development*. Model ini merupakan suatu model yang di dalamnya merepresentasikan tahapan-tahapan secara sistematis (tertata) dan sistemis dalam penggunaan bertujuan untuk tercapainya hasil yang di inginkan.

Analysis (Analisis)

Pada tahap ini, analisis dilakukan untuk menentukan tujuan dan batasan materi yang akan dipelajari. Tujuan lainnya adalah untuk mengumpulkan informasi tentang masalah atau hambatan yang muncul dalam proses pembelajaran serta tentang cara menggunakan sumber belajar dalam proses pembelajaran. Pada tahap ini dilakukan 2 tahap analisis yaitu 1). Analisis kinerja mencakup metode yang digunakan pendidik, sumber belajar yang digunakan di sekolah serta instrumen penilaian, pencapaian kemampuan berpikir kritis, dan hambatan yang dihadapi peserta didik selama proses pembelajaran; 2). Analisis kebutuhan yang mencakup analisis, kurikulum, analisis materi dan analisis karakteristik peserta didik.

Design (Perancangan)

Tahap desain atau biasa disebut dengan perancangan merupakan strategi dalam merancang dan mengembangkan program dengan mengidentifikasi bagian-bagian yang diperlukan pada proses pengembangan.

a. Penetapan Tujuan Tes

Tujuan penelitian ini berupa tes keterampilan berpikir kritis yang bertujuan untuk melihat kelemahan-kelemahan peserta didik serta faktor penyebabnya

b. Kisi-kisi tes

Kisi-kisi tes digunakan untuk menjadi pedoman dalam menyusun instrumen soal sesuai dengan kompetensi dasar yang digunakan dalam penelitian. Adapun bagian-bagian dalam kisi-kisi tes adalah capaian pembelajaran.

Development (Pengembangan)

Pada tahapan pengembangan, dilakukan pengoreksian kesalahan ataupun kekurangan pada instrumen tes agar sesuai dengan tujuan pembelajaran yang ingin dicapai. Pengoreksian ini dilakukan oleh tenaga ahli, tenaga ahli memberikan saran dan memvalidasi instrumen soal agar mengetahui adanya ketidaksesuaian instrumen soal yang disusun. Proses ini juga menentukan kelayakan instrumen tes sebelum diimplementasikan.

Penulisan butir soal berdasarkan pengembangan indikator berupa keterampilan berpikir kritis bentuk pertanyaan yang disesuaikan dengan kompetensi dasar, kisi-kisi tes, pokok bahasan, penomoran soal dan kunci jawaban. Dalam penelitian ini, terdapat 10 pertanyaan berupa soal uraian (*essay*) yang dikembangkan dari indikator yang digunakan. Penulisan soal menggunakan bahasa yang mudah dipahami dengan aktivitas sehari-hari peserta didik,

sehingga meminimalisir kesalahan dalam mentafsirkan pertanyaan.

Selanjutnya instrumen yang dibuat berupa angket validasi instrumen dan instrument tes keterampilan berpikir kritis. Tahapan validasi instrumen dilakukan dengan memberikan instrumen validasi yang terdiri dari 9 soal yang valid, indikator soal dan kunci jawaban. Selain itu dalam tahapan validasi, ahli menilai dengan pilihan esensial, tidak esensial dan tidak relevan. Ahli juga dapat memberikan saran dan kritik terhadap masing-masing butir soal ataupun secara keseluruhan. Hingga sampai penilaian akhir ahli dapat menentukan instrumen tes tersebut layak digunakan atau tidak layak. Pada langkah ini dilakukan uji validitas peserta didik (uji isi). Uji coba terbatas dilakukan pada 30 peserta didik yang sudah belajar materi ekologi dan keanekaragaman hayati indonesia.

Teknik Analisis Data

Data hasil uji coba instrumen dianalisis menggunakan analisis butir soal yang meliputi uji validitas, tingkat kesukaran, dan daya pembeda. Uji validitas butir dilakukan untuk mengetahui ketepatan setiap butir soal dalam mengukur aspek yang dimaksud. Analisis tingkat kesukaran dilakukan untuk mengelompokkan soal ke dalam kategori mudah, sedang, dan sukar, sedangkan daya pembeda digunakan untuk mengetahui kemampuan soal dalam membedakan peserta didik berkemampuan tinggi dan rendah. Berdasarkan hasil analisis, diperoleh 9 butir soal yang memenuhi kriteria kelayakan dan dinyatakan valid.

Hasil dan Pembahasan

Penelitian ini menghasilkan instrumen keterampilan berpikir kritis peserta didik materi ekologi dan keanekaragaman hayati indonesia kelas VII. Instrumen yang dikembangkan terlebih dahulu telah di validasi oleh validator ahli.

Pada tahap pertama yaitu tahap *analysis* (analisis), analisis awal diperlukan untuk memperoleh gambaran tentang pengembangan instrumen. Analisis tersebut mencakup; (a) analisis kinerja, (b) analisis kebutuhan. Adapun uraian analisis adalah sebagai berikut:

Analisis Kinerja

Tahap pertama dalam penelitian ini adalah analisis kinerja. Analisis kinerja didapatkan berdasarkan dari hasil observasi yang telah dilakukan, yaitu di SMP Negeri 27 Makassar. Berdasarkan pengamatan, proses pembelajaran yang diterapkan oleh pendidik menggunakan model pembelajaran *discovery learning*, yang tampak dari aktivitas guru mendorong peserta

didik untuk menemukan konsep secara mandiri melalui keterlibatan aktif dalam kegiatan pembelajaran, seperti mengamati, mengajukan pertanyaan, mengumpulkan informasi, serta menarik kesimpulan melalui diskusi dan presentasi di kelas. Dalam pelaksanaan pembelajaran, guru juga memanfaatkan beberapa sumber belajar seperti buku paket, power point serta sumber belajar digital lainnya. Namun, dari hasil observasi bahwa penilaian yang digunakan menggunakan penilaian tes biasa, belum pernah menggunakan tes keterampilan berpikir kritis.

a. Analisis kinerja

Analisis ini mencakup aspek metode pembelajaran yang digunakan, sumber dan media pembelajaran, serta instrumen penilaian yang diterapkan di sekolah, termasuk kendala yang dihadapi selama kegiatan pembelajaran. Berdasarkan hasil observasi dan analisis kinerja di SMP 27 Makassar, diketahui bahwa instrumen penilaian sebenarnya telah tersedia, namun belum mengakomodasi penilaian terhadap keterampilan berpikir kritis peserta didik. Kondisi ini menunjukkan bahwa belum optimalnya penggunaan instrumen yang mengukur keterampilan berpikir kritis dapat menghambat proses evaluasi dan refleksi pembelajaran secara lebih mendalam.

b. Analisis Kebutuhan

Analisis kebutuhan dilakukan dengan meninjau tiga aspek utama, yaitu analisis kurikulum, analisis materi, dan analisis karakteristik peserta didik. Ketiga aspek ini diperoleh melalui kegiatan observasi langsung dan wawancara dengan pendidik IPA di SMP Negeri 27 Makassar.

1) Analisis Kurikulum

Berdasarkan hasil observasi dan wawancara dengan pendidik di SMP Negeri 27 Makassar, pembelajaran IPA telah menerapkan Kurikulum Merdeka yang menekankan pengembangan kompetensi peserta didik secara holistik, termasuk kemampuan berpikir kritis sebagai bagian dari kemampuan berpikir tingkat tinggi. Namun, instrumen penilaian yang digunakan masih berupa tes konvensional yang belum dirancang untuk mengukur kemampuan berpikir kritis secara spesifik. Soal-soal yang digunakan cenderung berada pada level kognitif rendah hingga menengah sehingga belum mampu mengungkap kemampuan berpikir tingkat tinggi peserta didik secara optimal. Kondisi tersebut menunjukkan adanya kesenjangan antara tuntutan Kurikulum Merdeka yang berorientasi pada pengembangan kemampuan berpikir tingkat tinggi dengan instrumen penilaian yang digunakan dalam proses pembelajaran.

2) Analisis Materi

Analisis materi dilakukan dengan mengidentifikasi materi yang relevan untuk dijadikan konteks pengukuran kemampuan berpikir kritis. Materi yang dipilih adalah Ekologi dan Keanekaragaman Hayati Indonesia, yang merupakan salah satu materi pokok IPA kelas VII dalam Kurikulum Merdeka. Pemilihan materi ini didasarkan pada beberapa pertimbangan berikut.

Pertama, materi ekologi dan keanekaragaman hayati memiliki karakteristik yang kaya akan fenomena nyata dan isu lingkungan kontekstual, seperti pencemaran ekosistem, ancaman terhadap keanekaragaman hayati, dan interaksi antarorganisme. Fenomena-fenomena ini sangat potensial untuk memunculkan pertanyaan yang menuntut kemampuan analisis, evaluasi, dan refleksi mendalam dari peserta didik. Kedua, materi ini memiliki keterkaitan langsung dengan kehidupan sehari-hari peserta didik, sehingga relevan untuk mendorong proses berpikir reflektif yang bersifat personal dan kontekstual. Ketiga, kompleksitas konsep dalam materi ini mulai dari tingkatan ekosistem, interaksi biotik-abiotik, hingga upaya konservasi memberikan ruang yang luas untuk mengembangkan soal-soal berpikir kritis yang menuntut perbandingan, inferensi, dan evaluasi berdasarkan bukti.

3) Analisis Karakteristik Peserta Didik

Berdasarkan hasil observasi selama proses pembelajaran di kelas VII SMP Negeri 27 Makassar, ditemukan sejumlah kondisi yang menggambarkan karakteristik peserta didik. Pertama, sebagian peserta didik mengalami kesulitan dalam mempertahankan konsentrasi selama pembelajaran berlangsung, terutama ketika metode yang digunakan bersifat monoton. Kedua, peserta didik cenderung pasif dalam merespons pertanyaan yang membutuhkan penalaran tinggi dan lebih terbiasa menjawab pertanyaan faktual yang bersifat hafalan. Ketiga, ketika diberikan soal yang menuntut analisis atau refleksi, sebagian besar peserta didik kesulitan mengorganisasi jawaban secara sistematis dan mendalam.

Kondisi ini mengindikasikan bahwa kemampuan berpikir kritis dan berpikir reflektif peserta didik kelas VII di SMP Negeri 27 Makassar belum berkembang secara optimal. Salah satu faktor yang berkontribusi terhadap kondisi ini adalah ketiadaan instrumen penilaian yang secara khusus dirancang untuk mengukur kedua kemampuan tersebut. Oleh karena itu, pengembangan instrumen tes kemampuan berpikir kritis pada materi ekologi dan keanekaragaman hayati Indonesia menjadi kebutuhan yang mendesak dan relevan untuk dipenuhi.

Tahap kedua adalah tahap *desig* (desain). Tahapan ini mendesain atau merancang instrumen yang

merupakan salah satu tahapan penting dalam mengembangkan instrumen penelitian ini. Tahap desain yang dilakukan yaitu merancang instrumen keterampilan berpikir kritis dan kemampuan berpikir reflektif. Tahap ini dilakukan secara sistematis yang dimulai dengan pengumpulan data yang diperlukan dalam pengembangan instrumen. Dalam tahap ini menentukan materi yaitu ekologi dan keanekaragaman hayati Indonesia yang akan dimasukkan dalam pengembangan ini. Dimulai dengan menyusun kisi-kisi soal. Penyusunan kisi-kisi soal disesuaikan dengan indikator keterampilan berpikir kritis dapat dilihat pada tabel berikut.

Tabel 1. Kisi-kisi instrumen

Indikator	Sub Indikator	No butir soal	Jumlah Soal
Memberikan penjelasan sederhana (<i>elementary clarification</i>)	Memfokuskan pertanyaan	1,2,3	3
	Menganalisis argumen	4,5	2
Membangun keterampilan dasar (<i>basic support</i>)	Mempertimbangkan kredibilitas suatu sumber	6,7	2
	Mengobservasi dan mempertimbangkan hasil observasi	8,9,10	3

Tahap ketiga adalah *development* (pengembangan) dimana proses pencetakan produk yang kemudian siap untuk dilakukan proses validasi. Uji validasi dilakukan kepada 2 validator ahli dengan menggunakan lembar validasi yang telah disiapkan, validasi dilakukan guna memastikan butir soal-soal yang telah dibuat sesuai dengan indikator keterampilan berpikir keiti. Pengembangan instrumen ini mencakup penyusunan soal uraian sebanyak 10 butir soal pada materi ekologi dan keanekaragaman hayati Indonesia untuk peserta didik kelas VII. Setiap butir soal dirancang agar sesuai dengan indikator yang diukur serta mempertimbangkan tingkat kesulitan dan keterbacaan bagi peserta didik. Sebelum diujicobakan, instrumen divalidasi lebih dari tiga kali oleh masing-masing validator dengan sejumlah saran perbaikan yang mencakup aspek materi, konstruk soal, dan bahasa.

Hasil validasi ahli dirangkum pada tabel di bawah ini:

Tabel 2. Rangkuman Hasil Validasi

No	Aspek Penilaian	Rata-rata skor V1	Rata-rata skor V2	Content Validity	Ket
1	Kesesuaian isi	3,75	3,25	1	Valid
2	Konstruksi soal	4	3,5	1	Valid
3	Bahasa yang digunakan	3,5	3,5	1	Valid
	Rata-Rata	3,75	3,42	1	Valid

Nilai rata-rata total content validity yaitu 1 yang dinyatakan dalam kategori valid. Dinyatakan valid apabila nilai hasil content validity berada pada nilai $\geq 0,75$.

Setelah instrumen dinyatakan valid oleh validator ahli, instrumen diujicobakan kepada 30 peserta didik kelas VII. Butir soal dinyatakan valid apabila nilai r hitung $> r$ tabel (0,361) pada taraf signifikansi 5%.

Tabel 3. Data Hasil Uji Perbutir Soal

No. Soal	r Hitung	Validitas	Tingkat Kesukaran	Ket. TK	Daya Pembeda	Ket. DP	Simpulan
1	0,638	Valid	0,48	Sedang	0,30	Cukup	Dipakai
2	0,729	Valid	0,44	Sedang	0,30	Cukup	Dipakai
3	0,039	Tidak Valid	0,83	Mudah	-0,03	Buruk Sekali	Direvisi
4	0,686	Valid	0,52	Sedang	0,31	Cukup	Dipakai
5	0,601	Valid	0,66	Sedang	0,27	Cukup	Dipakai
6	0,442	Valid	0,78	Mudah	0,24	Cukup	Dipakai
7	0,376	Valid	0,69	Sedang	0,19	Buruk	Dipakai
8	0,693	Valid	0,75	Mudah	0,29	Cukup	Dipakai
9	0,697	Valid	0,61	Sedang	0,47	Baik	Dipakai
10	0,701	Valid	0,67	Sedang	0,46	Baik	Dipakai
Jumlah Valid		9 soal (1,2,4,5,6,7,8,9, dan 10)					
Jumlah Tidak Valid		1 soal (3)					

Analisis butir soal merupakan suatu prosedur sistematis yang memberikan informasi khusus terhadap butir soal yang disusun untuk memperoleh gambaran kualitas soal yang sebenarnya. Analisis ini penting dilakukan untuk mengetahui apakah setiap item soal benar-benar baik dan layak digunakan sebagai alat evaluasi. Berikut pembahasan rinci per butir soal berdasarkan analisis yang telah dilakukan:

Soal 1 dengan kategori valid, soal ini memiliki validitas yang baik karena r hitung (0,638) $> r$ tabel (0,361), menunjukkan soal mampu mengukur aspek keterampilan berpikir kritis yang dimaksud. Tingkat kesukaran 0,48 termasuk kategori sedang artinya soal tidak terlalu mudah dan tidak terlalu sulit sehingga ideal untuk digunakan. Daya pembeda 0,30 termasuk kategori cukup yang menunjukkan soal cukup mampu membedakan peserta didik kelompok atas dan bawah. Soal ini layak digunakan.

Soal 2 dengan kategori valid, soal ini memiliki validitas tertinggi di antara semua soal dengan (0,79), menunjukkan konsistensi yang sangat baik dalam mengukur indikator berpikir kritis. Tingkat kesukaran 0,44 tergolong sedang, proporsi jawaban benar cukup seimbang. Daya pembeda 0,30 termasuk kategori cukup, artinya soal ini sangat efektif dalam mengidentifikasi

peserta didik berprestasi dan yang memerlukan perhatian lebih. Soal ini direkomendasikan untuk tetap digunakan.

Soal 3 dengan kategori tidak valid, soal ini memiliki korelasi negatif yang mengindikasikan peserta didik dengan skor total tinggi cenderung menjawab salah, sedangkan skor rendah menjawab benar. Hal ini menunjukkan soal tidak berfungsi sebagaimana mestinya. Tingkat kesukaran 0,83 berada pada kategori mudah, yang berarti banyak peserta didik dapat menjawab benar. Daya pembeda negatif -0,03 termasuk kategori buruk sekali, karena soal lebih banyak dijawab benar oleh kelompok bawah. Kelemahan ini dapat disebabkan oleh redaksi soal yang ambigu, kunci jawaban yang kurang tepat, atau materi yang belum dipahami secara merata. Soal ini harus direvisi total atau diganti.

Soal 4 dengan kategori valid, soal ini memiliki validitas sangat baik dengan (0,686) dengan tingkat kesukaran sedang (0,52) yang ideal. Daya pembeda 0,31 termasuk kategori cukup, menunjukkan soal ini memiliki kemampuan yang baik dalam membedakan peserta didik kelompok atas dan bawah. Soal ini layak digunakan.

Soal 5 dengan kategori valid, soal ini memiliki validitas yang baik (0,601) dengan tingkat kesukaran sedang (0,66). Daya pembeda (0,27) dengan kategori cukup, sehingga soal ini sudah cukup baik dalam membedakan kemampuan peserta didik.

Soal 6 dengan kevalidan yang baik (0,442) yang menunjukkan bahwa butir soal memiliki hubungan yang memadai dengan skor total tes dapat digunakan untuk mengukur kompetensi yang menjadi tujuan instrumen. Tingkat kesukaran mudah (0,78), yang artinya peserta didik dapat menjawab soal dengan benar. Daya pembeda berada pada kategori cukup (0,24), yang menunjukkan bahwa soal masih memiliki kemampuan untuk membedakan peserta didik berkemampuan tinggi dan rendah.

Soal 7 dengan kevalidan (0,376), yang menunjukkan bahwa butir soal masih memiliki hubungan yang memadai dengan skor total tes dan dapat digunakan untuk mengukur kompetensi yang relevan dengan tujuan instrumen. Namun, nilai daya pembeda sebesar 0,19 berada pada kategori buruk, yang menunjukkan bahwa kemampuan soal dalam membedakan peserta didik berkemampuan tinggi dan rendah masih rendah. Secara teoritis, daya pembeda yang rendah dapat disebabkan oleh berbagai faktor, seperti soal yang terlalu mudah, pilihan jawaban yang kurang berfungsi sebagai pengecoh, materi yang belum dikuasai secara merata, atau adanya permasalahan dalam konstruksi soal. Meskipun demikian, berdasarkan hasil analisis pada tabel, butir soal nomor 7 dinyatakan dipakai karena validitasnya terpenuhi dan tingkat kesukarannya berada pada kategori sedang, tetapi soal tersebut tetap perlu dievaluasi dan diperbaiki agar kemampuan daya pembedanya dapat meningkat pada penggunaan instrumen berikutnya.

Soal nomor 8 memperoleh nilai r hitung sebesar 0,693 dan dinyatakan valid, yang menunjukkan adanya hubungan yang kuat antara skor pada butir soal dengan skor total tes sehingga soal mampu mengukur kompetensi yang sesuai dengan tujuan instrumen. Nilai tingkat kesukaran sebesar 0,75 berada pada kategori mudah, yang berarti sebagian besar peserta didik mampu menjawab soal dengan benar. Soal yang mudah dapat berfungsi untuk mengetahui penguasaan konsep dasar peserta didik, tetapi apabila digunakan dalam jumlah yang terlalu banyak dapat mengurangi kemampuan instrumen dalam membedakan tingkat kemampuan peserta didik. Nilai daya pembeda sebesar 0,29 termasuk kategori cukup, yang menunjukkan bahwa soal masih mampu membedakan peserta didik berkemampuan tinggi dan rendah dalam tingkat yang cukup. Berdasarkan hasil tersebut, butir soal nomor 8 dinyatakan dipakai karena memiliki validitas yang tinggi dan daya pembeda yang cukup, meskipun tingkat

kesukarannya perlu diperhatikan dalam penyusunan instrumen selanjutnya.

Soal nomor 9 memperoleh nilai r hitung sebesar 0,697 dan dinyatakan valid, yang menunjukkan bahwa terdapat hubungan yang kuat antara kemampuan peserta didik dalam menjawab butir soal dengan skor total tes. Nilai tingkat kesukaran sebesar 0,61 termasuk kategori sedang, sehingga soal memiliki tingkat kesulitan yang proporsional dan mampu memberikan kesempatan kepada peserta didik dengan berbagai tingkat kemampuan untuk menunjukkan penguasaannya. Selain itu, nilai daya pembeda sebesar 0,47 berada pada kategori baik, yang menunjukkan bahwa soal mampu membedakan peserta didik yang memiliki kemampuan tinggi dengan peserta didik yang memiliki kemampuan rendah secara efektif. Berdasarkan ketiga indikator tersebut, butir soal nomor 9 dinyatakan dipakai dan dapat dipertahankan dalam instrumen karena memiliki validitas yang tinggi, tingkat kesukaran sedang, serta daya pembeda yang baik.

soal nomor 10 memperoleh nilai r hitung sebesar 0,701 dan dinyatakan valid, yang menunjukkan bahwa butir soal memiliki hubungan yang kuat dengan skor total tes dan mampu mengukur kompetensi yang sesuai dengan tujuan instrumen. Nilai tingkat kesukaran sebesar 0,67 termasuk kategori sedang, yang menunjukkan bahwa soal memiliki tingkat kesulitan yang relatif proporsional dan tidak terlalu mudah maupun terlalu sulit. Sementara itu, nilai daya pembeda sebesar 0,46 termasuk kategori baik, yang berarti soal mampu membedakan peserta didik dengan kemampuan tinggi dan rendah secara efektif. Peserta didik yang memiliki penguasaan kompetensi lebih baik cenderung mampu menjawab soal dengan benar dibandingkan peserta didik yang memiliki penguasaan kompetensi lebih rendah. Berdasarkan hasil analisis tersebut, butir soal nomor 10 dinyatakan dipakai dan dapat dipertahankan karena memiliki validitas yang baik, tingkat kesukaran sedang, serta daya pembeda yang baik.

Secara keseluruhan, dari 10 butir soal yang diuji coba, diperoleh 9 soal valid (nomor 1, 2, 4, 5, 6, 7, 8, 9, dan 10) serta 1 soal tidak valid (nomor 3). Hasil ini sejalan dengan teori Fatih (2026) analisis butir soal yang menegaskan pentingnya validitas, daya pembeda, dan indeks kesukaran sebagai indikator utama kelayakan instrumen evaluasi.

Sejalan dengan Sudrajat (2024), validitas, tingkat kesukaran, dan daya pembeda merupakan formula penting dalam menyeleksi butir soal. Soal yang baik tidak boleh terlalu sukar atau terlalu mudah serta harus mampu membedakan kemampuan siswa secara tajam. Penggunaan instrumen berbasis keterampilan berpikir kritis ini juga relevan dengan tuntutan

Kurikulum Merdeka dan pembelajaran abad ke-21 yang berfokus pada pengembangan kemampuan berpikir tingkat tinggi (Higher Order Thinking Skills).

Kesimpulan

Berdasarkan hasil dan pembahasan, pengembangan instrumen tes keterampilan berpikir kritis pada materi Ekologi dan Keanekaragaman Hayati Indonesia untuk peserta didik kelas VII SMP dinyatakan valid dan layak digunakan setelah melalui tahap uji ahli. Hasil validasi isi oleh ahli menunjukkan skor *content validity* sempurna sebesar 1 (kategori valid), sementara uji coba empiris terhadap 10 butir soal uraian menghasilkan 9 soal valid dan layak dipakai (nomor 1, 2, 4, 5, 6, 7, 8, 9, dan 10) karena memiliki tingkat kesukaran yang seimbang (sedang hingga mudah) serta daya pembeda yang memadai (cukup hingga baik). Adapun 1 butir soal dinyatakan tidak valid (nomor 3) akibat nilai korelasi dan daya pembeda yang buruk, sehingga memerlukan revisi total atau penggantian sebelum dapat dimanfaatkan secara optimal.

Daftar Pustaka

- Aziz, V. M. S., & Adi, S. A. (2024). Upaya Mengembangkan Kemampuan Berpikir Kritis melalui Challenge Based on Ethnomathematics Learning berbantuan Wordwall dan AI-Video. *Jurnal Ilmiah Mahasiswa Pendidikan Matematika*. 4(1). 151-160.
- Darlis, A., Ali, I. S., Musthafa, F. P., Lisa, S., & Isnayani, R. (2022). Pendidikan Berbasis Merdeka Belajar. *Jurnal Ilmu-ilmu Keislaman*. 11(2). 393-401.
- Faith, N. I., Ma'mun, H., & Dewi, M. (2026). Konsep dan Implementasi Analisis Butir Soal Dalam Evaluasi Pembelajaran. *Jurnal Ilmiah Pendidikan Dasar*. 11(2). 231-244
- Kurniasari, I., & Hanin. N. F. (2022). Model Pembelajaran Creative Problem Solving (CPS) Berbasis Socioscientific untuk Meningkatkan Kemampuan Berpikir Reflektif Peserta Didik. *Jurnal Tadris IPA Indonesia*. 2(3). 272-282.
- Kurniawati, E. F., & Yuyu, Y. (2024). Pengembangan Instrumen Tes Untuk Kemampuan Berpikir Reflektif Matematis Pada Materi Aljabar. *Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*. 5(2). 1021-1030.
- Marfhadella, P. (2021). Pengembangan Assessment untuk Mengukur Kemampuan Berpikir Kritis Materi Interaksi Makhluk Hidup dengan Lingkungan di SMP se kota Bengkulu. Diperoleh dari <http://repository.iainbengkulu.ac.id/7156/1/Skripsi%20Putri%20Marfhadella.pdf>
- Maulina, N., Ramlawati., & Salma, S. (2026). Analisis Kemampuan Berpikir Tingkat Tinggi Peserta Didik pada Pembelajaran Pesawat Sederhana Di Sekolah Menengah Pertama. *Pinisi Journal Of Education*. 6(3). 150-167.
- Mudakir, N. F., Joko, S., & Nurma, A. (2020). Kemampuan Berpikir Reflektif Matematis Siswa Dalam Menyelesaikan Soal Program Linear. *Jurnal Pendidikan MIPA*. 5(1). 21-26.
- Putri, S. M., Fitri A., Muhyiatul, F., & Ria, A. (2023). Validitas Instrumen Soal Tes Keterampilan Berpikir Kritis Siswa Kelas X pada Materi Komponen Ekosistem dan Interaksinya. *Jurnal Pendidikan Tambusai*. 7(3). 24253-24261.
- Rifda, N. H. A. (2023). Analisis Kemampuan Berpikir Peserta Didik. *Jambura Journal of Educational Management*. 4(2). 388-394.
- Sudrajat. (2024). Analisis Kualitas Instrumen Tes Untuk Mengukur Kemampuan Komunikasi Matematika Siswa SMP. *Jurnal Pendidikan Matematika*. 5(2). 210-221
- Tuerah, R. M., & Jeanne, M. T. (2023). Kurikulum Merdeka dalam Perspektif Kajian Teori: Analisis Kebijakan untuk Peningkatan Kualitas Pembelajaran di Sekolah. *Jurnal Ilmiah Wahana Pendidikan*. 9(19). 979-988.