

Pengembangan Instrumen Tes Essay Berbasis Kemampuan Berpikir Kritis Materi Suhu dan Kalor Pada Pembelajaran IPA di Sekolah Menengah Pertama

Nurul Nabilah ✉, Diva Khairina Umma, Sova Marwah Nasution, Mentari Darma Putri

Universitas Samudra

Jl. Prof. Dr. Syarief Thayeb, Meurandeh, Langsa Lama, Kota Langsa, Aceh 24416, Indonesia

| nrlnabila89@gmail.com ✉ | DOI : <https://doi.org/10.37729/jips.v6i2.5999> |

Article Info

Submitted

10/01/2025

Revised

14/10/2025

Accepted

10/11/2025

Abstrak – Kemampuan berpikir merupakan proses berpikir menganalisis, mengevaluasi, dan mensintesis informasi secara logis. Keterampilan ini berperan penting dalam membantu siswa memperoleh pemahaman yang lebih mendalam terhadap materi pelajaran. Tujuan dari penelitian ini adalah untuk mengembangkan alat ujian esai berbasis keterampilan berpikir kritis pada materi suhu dan kalor untuk siswa sekolah menengah di Langsa provinsi Aceh. Penelitian ini mengadopsi metode penelitian dan pengembangan (R&D) berdasarkan model pengembangan ADDIE, yang terdiri dari lima tahap: analisis, desain, pengembangan, implementasi, dan evaluasi. Penelitian ini melibatkan 27 siswa kelas VII dari salah satu sekolah menengah di kota Langsa. Pengumpulan data dilakukan melalui tes esai 10 pertanyaan untuk mengukur keterampilan berpikir kritis siswa. Hasil uji instrumen dianalisis menggunakan aplikasi ANATES untuk mengetahui validitas, reliabilitas, daya pembeda, dan tingkat kesulitan soal. Analisis data menunjukkan bahwa dari 10 pertanyaan, 9 pertanyaan dinyatakan valid dan 1 pertanyaan dinyatakan tidak valid. Nilai reliabilitas sebesar 0,94 menunjukkan konsistensi yang sangat baik. Analisis spesifisitas menghasilkan sembilan pertanyaan yang diklasifikasikan dalam kategori “sangat baik” dan satu pertanyaan yang diklasifikasikan dalam kategori “diragukan”. Dilihat dari tingkat kesukaran soal, terdapat 7 soal yang tergolong sedang dan 3 soal yang tergolong sukar. Secara keseluruhan instrumen tes yang dikembangkan terbukti valid dan reliabel dalam mengukur kemampuan berpikir kritis siswa.

Kata kunci: Berpikir kritis, Instrumen tes, Tes essay

Abstract – Critical thinking skills are the thinking process of analyzing, evaluating and synthesizing information logically. This skill plays an important role in helping students gain a deeper understanding of the subject matter. The aim of this research is to develop an essay test tool based on critical thinking skills on temperature and heat material for middle school students in Langsa, Aceh province. This research adopts a research and development (R&D) method based on the ADDIE development model, which consists of five stages: analysis, design, development, implementation, and evaluation. This research involved 27 class VII students from one of the middle schools in Langsa city. Data collection was carried out through a 10 question essay test to measure students' critical thinking skills. The results of the instrumental experiment were analyzed using the ANATES application to determine the validity, reliability, distinguishing power and level of difficulty of the questions. Data analysis shows that of the 10 questions, 9 questions were declared valid and 1 question was declared invalid. The reliability value of 0.94 shows very good consistency. Specificity analysis resulted in nine questions being classified in the “very good” category and one question being classified in the “doubtful” category. Judging from the level of difficulty of the questions, there are 7 questions that are classified as medium and 3 questions that are classified as difficult. Overall, the test instrument developed was proven to be valid and reliable in measuring students' critical thinking abilities



Keywords: Critical thinking, Instrument test, Essay test

1. Pendahuluan

Pendidikan merupakan faktor yang sangat penting bagi pembangunan suatu negara. Kualitas pendidikan, baik atau buruk, menunjukkan keberhasilan pemerintahan suatu negara, selain itu turut dipengaruhi pula oleh faktor seperti kurikulum, guru, fasilitas, dan sumber belajar.

Guru mempunyai peranan yang sangat penting dalam meningkatkan kualitas pembelajaran di kelas. Dengan menggunakan pendekatan yang menempatkan siswa sebagai pusat pembelajaran, guru dapat meningkatkan proses pembelajaran. Namun kemajuan pendidikan Indonesia terhambat karena banyak guru yang terus menggunakan pendekatan pembelajaran yang monoton dan tidak efektif untuk meningkatkan kemampuan berpikir kritis siswa [1], [2]. Rendahnya kemampuan berpikir kritis siswa Indonesia disebabkan oleh tidak efisiennya proses pembelajaran yang tidak sesuai dengan minat siswa [3], [4]. Guru mempunyai pengaruh yang cukup besar terhadap proses pendidikan, hal tersebut berkaitan dengan sangat pentingnya peran guru yang bertanggung jawab terhadap keberhasilan pendidikan. Menurut data studi *The Trends in International Mathematics and Science Study* lebih dari 95% siswa Indonesia hanya dapat menjawab soal dengan tingkat kesulitan sedang, sedangkan sekitar 50% siswa Taiwan hanya dapat menjawab soal dengan tingkat kesulitan tinggi. Faktor utamanya adalah siswa Indonesia kurang memiliki kemampuan dalam memecahkan masalah situasional yang memerlukan kreativitas, pemikiran logis, dan penalaran [5], [6], [7].

Sebagian siswa memiliki pola berpikir pasif dengan menghafal persamaan tanpa memahami penerapannya. Memberikan pertanyaan berpikir kritis memungkinkan siswa mempertanyakan asumsi, menggali lebih dalam konten, dan menghubungkan konsep yang dipelajari dengan situasi dunia nyata. Pertanyaan-pertanyaan dalam soal tersebut mendorong siswa untuk berpikir kreatif dan inovatif, dengan mencermati ragam masalah dari berbagai sudut pandang. Dengan cara tersebut, siswa tidak hanya memperoleh pemahaman lebih dalam tentang mata pelajaran tetapi juga mengembangkan sikap positif untuk menghadapi tantangan. Menurut penelitian Jiwandono [8] alat penilaian kurang cocok untuk melatih keterampilan berpikir kritis karena isi soal tidak memuat indikator berpikir kritis. Selain itu, penelitian ini menemukan bahwa banyak siswa memberikan jawaban yang tidak tepat dan singkat, baik karena terlalu singkat atau karena tidak berpikir kritis. Hal ini karena hanya sedikit soal yang dirancang untuk mengajarkan keterampilan berpikir kritis kepada siswa dan tidak ada tes yang dirancang khusus untuk keterampilan berpikir kritis. Sementara itu, Nandar [9] menyatakan bahwa anak diajarkan berpikir kritis sejak dini agar mampu memikirkan masalah dan menyelesaikannya secara cermat.

Pertanyaan merupakan perangkat penilaian yang dirancang untuk mengetahui seberapa baik siswa memahami pelajaran dan keterampilan apa yang dimilikinya. Analisis unsur pertanyaan bertujuan untuk memperbaiki ujian dengan memodifikasi atau menghilangkan pertanyaan yang tidak efektif dan untuk menilai apakah siswa memahami isinya [10]. Dalam penelitian ini, kemampuan berpikir kritis siswa diuji dengan menggunakan soal essay. Pertanyaan-pertanyaan ini dirancang tidak hanya untuk mendorong siswa menjawab dengan benar, tetapi juga untuk menjelaskan bagaimana memahami dan menerapkan konsep suhu dan kalor. Tes essay merupakan soal yang dirancang dalam bentuk pertanyaan terstruktur dan mengharuskan peserta menuliskan jawabannya sendiri dalam bahasanya sendiri; dapat pula berbentuk tugas atau soal yang mengharuskan siswa mengorganisasikan konsep-konsep dan hal-hal yang telah dipelajarinya dalam tulisannya sendiri [11].

Menurut data PISA (*Program for International Student Assessment* 2018), siswa Indonesia memiliki kemampuan berpikir logis, khususnya berpikir kritis, dan mendapat peringkat pertama dari bawah, dan tingkat pencapaian berpikir kritis siswa Indonesia terbukti sangat rendah salah satunya pada prestasi literasi membaca (*reading literacy*), literasi matematika (*mathematical literacy*), dan literasi sains (*scientific literacy*). Hal ini menunjukkan bahwa kemampuan berpikir logis siswa Indonesia sangat rendah. Ada beberapa faktor yang menurunkan kemampuan berpikir kritis siswa. Salah satunya adalah siswa kecenderungan menghafal materi dan rumus dibandingkan memahami konsep. Akibatnya kemampuan berpikir kritis siswa tidak terlatih dengan baik dan menjadi bosan dengan konsep yang dipelajarinya [12], [13]. Adanya penurunan kemampuan siswa dalam berpikir kritis berdampak pada ketidakmampuan siswa dalam menganalisis dan memecahkan permasalahan nyata yang dihadapi dalam kehidupan sehari-hari sehingga menyulitkan dalam mengambil keputusan secara cepat dan tepat [4], [14].

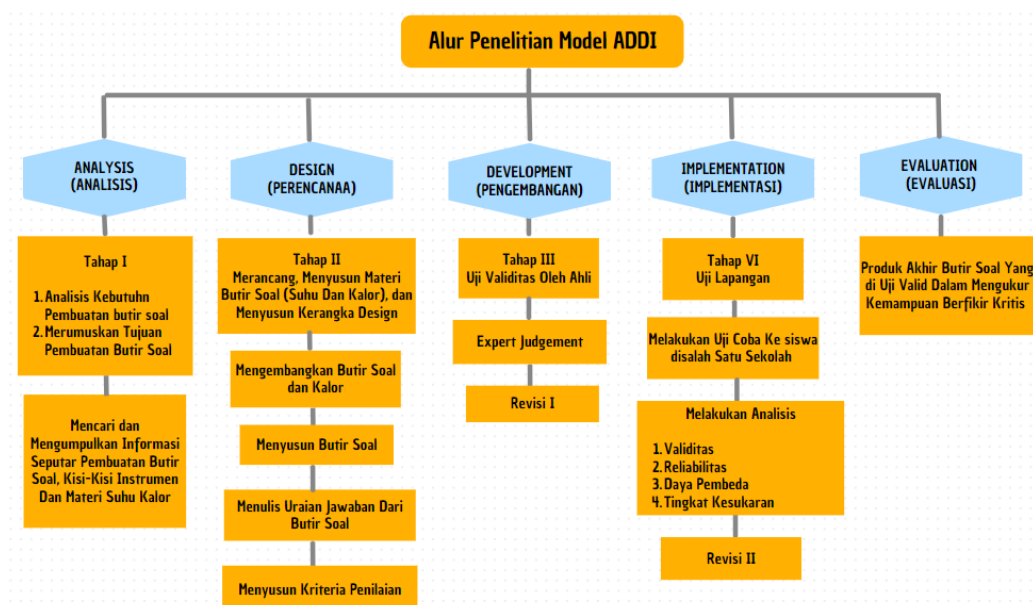
Mengingat pembahasan sebelumnya mengenai isu-isu seputar pemikiran kritis dan kurangnya minat di antara siswa dalam mengembangkan kemampuan berpikir kritis dan mengelola informasi, peneliti bermaksud untuk menilai butir-butir soal yang terkait dengan konsep kalor dan suhu.

Tujuannya adalah untuk memungkinkan siswa terlibat dalam pemikiran kritis, khususnya untuk menganalisis (C-4) dan mengevaluasi (C-5), sehingga mendorong mereka untuk secara aktif mencari dan mengelola informasi daripada menyerapnya secara pasif tanpa pemeriksaan apa pun. Investigasi terhadap butir-butir soal ini merupakan bagian dari inisiatif yang lebih luas untuk menawarkan pengalaman praktis kepada siswa dan pendidik masa depan saat mereka berupaya mengembangkan dan mengevaluasi kualitas butir-butir soal mengenai topik suhu dan kalor untuk siswa sekolah menengah pertama. Hal ini sejalan dengan penelitian [15] tentang pengembangan tes untuk mengukur HOTS, [16] tentang orientasi HOTS untuk meningkatkan berpikir kritis, dan [17] tentang bagaimana mengembangkan asesmen untuk mengukur HOTS. Materi ini dipilih karena merupakan bagian penting dari kurikulum fisika di tingkat SMP, dan pemahaman yang baik dari konsep ini merupakan landasan penting untuk pembelajaran fisika selanjutnya. Penelitian ini mengharapkan siswa dan calon pendidik dapat memperoleh pengalaman berharga dalam mengembangkan instrumen penilaian yang berkualitas, yang akan bermanfaat bagi mereka dalam menjalankan tugas sebagai pendidik di masa mendatang.

2. Metode

Jenis penelitian ini adalah penelitian pengembangan dengan menggunakan model pengembangan ADDIE untuk mengembangkan instrumen tes keterampilan berpikir kritis terkait suhu dan kalor. Kelas VII-1 Sekolah Menengah Pertama (SMP) di Kota Langsa merupakan lokasi penelitian yang dilakukan dengan jumlah siswa sebanyak 27 orang. ADDIE terdiri dari lima tahap: analisis, desain, pengembangan, implementasi, dan evaluasi. Berikut tabel yang akan menampilkan level ADDIE yang telah diterapkan. Adapun tahapan pengembangan dapat disajikan pada Gambar 1.

Data yang digunakan dalam penelitian ini meliputi data kualitatif dan kuantitatif. Data kuantitatif di peroleh melalui pengukuran dan perhitungan hasil pengerjaan soal siswa. Arikunto (2021) menyatakan bahwa data kuantitatif ialah data yang dapat dianalisis dengan menggunakan metode statistik. Dalam penelitian ini diperoleh data kuantitatif seperti validitas, reliabilitas, daya pembeda, dan tingkat kesukaran dari hasil tes instrumen soal. Selanjutnya penelitian ini dilakukan dengan menggunakan aplikasi ANATES untuk mendapatkan data yang lebih akurat. ANATES ialah perangkat lunak untuk menganalisis butir soal pada instrumen evaluasi seperti ulangan dan ujian. Aplikasi ini bertujuan agar pendidik dan peneliti dapat mengevaluasi kualitas soal tes berdasarkan karakteristik validitas, reliabilitas, daya pembeda, dan tingkat kesukaran



Gambar 1. Tahapan Pengembangan Instrumen Tes Dengan Model ADDIE

3. Hasil dan Pembahasan

Tujuan penelitian ini untuk menganalisis kualitas butir soal keterampilan berpikir kritis seperti validitas, reliabilitas, daya pembeda, dan tingkat kesukaran (indeks kesulitan) pada butir soal suhu dan kalor. Hasil penelitian ini diperoleh berdasarkan pengumpulan data yang dilakukan peneliti selama kegiatan pengujian butir soal yang dirancang dalam bentuk uraian atau essay. Aspek yang diamati adalah uji coba butir soal, kemampuan berpikir kritis siswa pada mata pelajaran IPA, suhu dan kalor. Soal esai kemudian dihitung menggunakan rumus validitas, reliabilitas, daya pembeda, dan tingkat kesukaran. Lalu dilanjutkan menggunakan ANATES untuk mengukur kualitas soal berdasarkan validitas, reliabilitas, daya pembeda dan tingkat kesukaran. hal ini dilakukan agar data yang didapat lebih akurat. maka diperoleh hasil sebagai berikut:

3.1. Validitas soal

Hasil perhitungan validitas instrumen tes dikumpulkan menggunakan aplikasi ANATES didapat sebanyak 9 soal dinyatakan valid, dan 1 soal dinyatakan cukup valid, berdasarkan korelasi skor dari 10 soal. Berikut adalah diagram tes hasil validitas. Untuk butir soal nomor 1 diperoleh indeks korelasi sebesar 0.857 pada kriteria valid, butir nomor 2 diperoleh 0.820 pada kriteria valid, soal nomor 3 diperoleh indeks korelasi sebesar 0.224 pada kriteria tidak valid, untuk soal nomor 4 termasuk kriteria signifikan/ valid dengan indeks korelasi sebesar 0.795. indeks korelasi berturut-turut untuk butir nomor 5 hingga 10 diperoleh: 0.715, 0.689, 0.853, 0.807, 0.689, dan 0.764 kesemuanya dalam kriteria valid. Adapun secara keseluruhan terdapat 90% butir soal yang valid dan 1 butir soal yang tidak valid. Butir-butir ini memiliki nilai korelasi di atas 0.600, menunjukkan bahwa soal tersebut berkorelasi tinggi dengan skor total tes. Ini berarti soal-soal tersebut dapat mengukur kemampuan yang seharusnya diukur. Butir nomor 3 memiliki korelasi $r = 0.224$, yang termasuk dalam kategori "tidak valid". Korelasi ini rendah dan menunjukkan bahwa butir soal tersebut kurang berkaitan dengan keseluruhan tes. Pada soal nomor 3 terjadi miskonsepsi. Adapun penyebab miskonsepsi seperti, kurangnya pemahaman tentang mekanisme perpindahan panas. butir soal yang tidak valid perlu diperbaiki untuk menghilangkan miskonsepsi siswa, sehingga seluruh soal dapat lebih efektif dalam mengukur pemahaman siswa secara ilmiah. Secara keseluruhan dapat disimpulkan jika mayoritas soal dalam tes memiliki validitas yang baik, menunjukkan bahwa soal-soal tersebut berkualitas. Namun, butir soal yang tidak valid perlu diperbaiki untuk menghilangkan miskonsepsi siswa, sehingga seluruh soal dapat lebih efektif dalam mengukur pemahaman siswa secara ilmiah.

3.2. Reliabilitas

Reliabilitas instrumen merupakan tingkat konsistensi yang ditunjukkan oleh instrumen ketika diberikan kepada subyek yang sama, terlepas dari variasi penyelenggara, waktu, atau lokasi. Menurut Sugiyono [18], Reliabilitas mencerminkan kemampuan instrumen untuk memberikan temuan yang serupa secara konsisten, yang sangat penting dalam konteks penelitian kemampuan berpikir kritis. Berikut adalah hasil korelasi reliabilitas. Hasil reliabilitas tes menunjukkan nilai 0.94, yang mengindikasikan bahwa tes memiliki tingkat reliabilitas tinggi. Artinya, instrumen pengukuran sangat konsisten dalam menilai kemampuan berpikir kritis siswa. Korelasi antara skor ganjil dan genap sebesar 0.88, yang menunjukkan hubungan yang kuat dan positif. Hal ini menegaskan bahwa item-item tes ganjil dan genap memiliki konsistensi yang baik. Rata-rata skor total siswa adalah 18.11, yang mengindikasikan bahwa rata-rata kemampuan berpikir kritis siswa berada di kategori menengah. Besar simpangan baku 8.79 menunjukkan adanya penyebaran skor yang cukup luas di antara siswa. Hal ini mengindikasikan perbedaan kemampuan berpikir kritis yang cukup signifikan di antara siswa yang terlibat dalam tes.

3.3. Daya Pembeda

Daya pembeda soal merupakan kemampuan suatu butir soal untuk membedakan siswa yang termasuk dalam kategori kelompok atas dan kelompok bawah berdasarkan kemampuan siswa. Butir soal yang memiliki daya pembeda tinggi akan lebih efektif dalam mengidentifikasi siswa yang memiliki pemahaman yang berbeda terhadap materi ujian. Rekapitulasi hasil perhitungan daya pembeda pada tes esai menggunakan aplikasi ANATES. Adapun data daya pembeda disajikan pada Tabel 1.

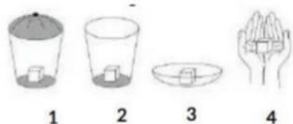
Tabel 1. Daya Pembeda

No	Daya Pembeda (%)	Kategori
1.	57.14	Sangat membeda
2.	53.57	Sangat membeda
3.	7.14	Butir diragukan
4.	46.43	Sangat membeda
5.	53.57	Sangat membeda
6.	46.43	Sangat membeda
7.	78.57	Sangat membeda
8.	42.86	Sangat membeda
9.	82.14	Sangat membeda
10.	67.86	Sangat membeda

Berdasarkan Tabel 1 menunjukkan bahwa dari 10 soal, 9 soal berada dalam kategori "sangat membeda"; hal tersebut menunjukkan bahwa sebagian besar soal sudah sesuai untuk digunakan sebagai alat evaluasi kemampuan siswa. Dan satu butir soal diragukan yang terdapat pada nomor 3. Pada soal nomor 3 merupakan soal pengecoh, dimana pada soal tersebut terdapat empat gambar dengan gambar pertama (gambar yang berbeda dengan gambar lainnya) yang mana gambar tersebut merupakan pilihan yang salah tetapi siswa malah memilih opsi/gambar tersebut. Jadi, pada soal tersebut terjadi miskonsepsi dimana siswa lebih banyak memilih gambar pertama (es batu di dalam gelas tertutup) sebagai jawaban yang paling cepat mencair dibandingkan gambar terakhir (es batu di telapak tangan). Miskonsepsi sama seperti ini pernah terjadi yaitu miskonsepsi tentang perubahan wujud es [19]. Adapun penyebab miskonsepsi lain seperti, kurangnya pemahaman tentang mekanisme perpindahan panas. Siswa belum memahami bahwa konduksi (kontak langsung dengan tangan) lebih efektif dalam mentransfer panas dibandingkan dengan transfer panas melalui udara dalam gelas. Pengetahuan yang belum terintegrasi, siswa salah mengaplikasikan konsep isolasi termal, berpikir bahwa gelas tertutup mempercepat pencairan karena "panas terjebak". Pengaruh pengalaman sehari-hari, siswa jarang mengamati fenomena langsung terkait mencairnya es batu dalam berbagai kondisi. Tak hanya itu, guru juga perlu menekankan pentingnya berpikir berdasarkan bukti (hasil percobaan) dan menghindari asumsi yang tidak didukung oleh pemahaman ilmiah. Dengan cara ini, miskonsepsi seperti ini dapat dikurangi. Solusi yang bisa dilakukan untuk soal nomor 3 agar tidak terjadi miskonsepsi adalah mengganti opsi pilihan yang tertera pada tabel 2.

Soal Sebelum Direvisi

Perhatikan gambar di bawah ini!

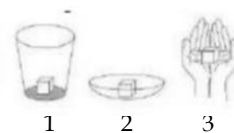


Gambar. 2 es batu dalam wadah

Pada setiap gambar di atas menunjukkan masing masing perbedaan, dimana pada setiap es batu ada pada wadah yang berbeda. Pada gambar pertama, es batu di letakkan di sebuah gelas dan ditutup, pada gambar kedua es batu diletakkan di sebuah gelas namun dalam keadaan gelas terbuka, gambar ketiga es batu diletakkan di wadah dalam kondisi terbuka, dan gambar terakhir es batu diletakkan di telapak tangan. Menurut pengamatanmu, es batu manakah yang paling cepat mencair, berikan alasanmu!

Soal Setelah Direvisi

Perhatikan gambar di bawah ini!



Gambar. 2 es batu dalam wadah

Pada setiap gambar di atas menunjukkan masing masing perbedaan, dimana pada setiap es batu ada pada wadah yang berbeda. Pada gambar pertama, es batu diletakkan di sebuah gelas dalam keadaan gelas tertutup, pada gambar kedua es batu diletakkan di wadah dalam kondisi terbuka, dan gambar terakhir es batu diletakkan di telapak tangan. Menurut pengamatanmu, es batu manakah yang paling cepat mencair, berikan alasanmu!

3.4. Tingkat Kesukaran

Tingkat kesukaran dapat diukur melalui indeks kesukaran, yaitu proporsi peserta yang berhasil menjawab soal dengan benar. Indeks ini tidak hanya memberikan gambaran tentang kesulitan soal, tetapi juga mengukur sejauh mana soal tersebut mampu merangsang kemampuan berpikir kritis siswa dalam menganalisis, mengevaluasi, dan menerapkan konsep untuk menemukan solusi yang tepat. Rekapitulasi hasil perhitungan tingkat kesukaran pada tes esai menggunakan aplikasi ANATES yaitu 70% soal dalam kategori sedang dan 30% dalam kategori sukar. Adapun rekapitulasi datanya dapat disajikan pada Tabel 2.

Tabel 2. Tingkat Kesukaran

No	Daya Pembeda (%)	Kategori
1.	53.57	Sedang
2.	33.93	Sukar
3.	50.00	Sedang
4.	37.50	Sukar
5.	44.64	Sedang
6.	44.64	Sedang
7.	50.00	Sedang
8.	39.29	Sukar
9.	51.79	Sedang
10.	51.79	Sedang

Berdasarkan Tabel 2 diatas menunjukkan dari 10 soal, 7 soal berada di kategori "Sedang," yang berarti soal tersebut memiliki tingkat kesukaran yang ideal. Namun, terdapat 3 soal dalam kategori "Sukar,". Tingkat kesukaran dengan kategori "Sedang," terdapat pada nomor 1, 3, 7, 9, dan 10 menunjukkan distribusi tingkat kesulitan yang baik dan sesuai untuk evaluasi pembelajaran. Soal nomor 5 dan 6 memiliki tingkat kesukaran yang mendekati batas bawah kategori sedang. Meski masih tergolong ideal, perlu perhatian lebih agar tetap relevan dan tidak bergeser ke kategori "Sukar." Soal dengan kategori "Sedang" sudah sangat baik untuk digunakan tanpa revisi signifikan. Soal ini dapat menjadi contoh untuk pembuatan soal di masa mendatang. Pada tingkat kesukaran dengan kategori "Sukar," terdapat pada nomor, 2, 4 dan 8 Soal Nomor menunjukkan bahwa siswa cenderung kesulitan dalam menjawab. Perlu evaluasi pada aspek kerumitan materi atau struktur soal. Soal nomor 4 mendekati batas bawah kategori "Sedang." Penyempurnaan kecil pada redaksi soal dapat membuatnya lebih mudah dipahami. Soal nomor 8 menunjukkan memiliki peluang untuk tetap relevan dengan sedikit perbaikan, terutama pada kejelasan konteks atau panduan. Soal kategori "Sukar" perlu dievaluasi dan direvisi. Fokus pada penyederhanaan redaksi, memberikan petunjuk yang lebih jelas, menyelaraskan soal dengan tingkat kemampuan siswa.

Instrumen tes essay berbasis kemampuan berpikir kritis pada materi suhu dan kalor yang dikembangkan dalam penelitian ini terdiri atas 10 butir soal yang dianalisis menggunakan ANATES. Hasil uji validitas menunjukkan sebagian besar soal memiliki validitas tinggi hingga sangat tinggi, sehingga dapat dinyatakan mampu mengukur konstruk yang diinginkan, yaitu kemampuan berpikir kritis siswa. Reliabilitas instrumen yang dihitung dengan koefisien Alpha Cronbach menunjukkan nilai tinggi, menandakan konsistensi internal yang baik. Analisis daya beda memperlihatkan bahwa mayoritas soal memiliki daya pembeda baik hingga sangat baik, meskipun terdapat beberapa soal yang perlu revisi. Tingkat kesukaran soal terdistribusi secara proporsional dengan dominasi kategori sedang, sehingga tes tidak terlalu mudah maupun terlalu sulit. Secara keseluruhan, instrumen ini layak digunakan sebagai alat evaluasi pembelajaran IPA di SMP.

Temuan ini sejalan dengan penelitian Lestari [20] yang mengembangkan instrumen tes berpikir kritis pada materi IPA SMP/MTs. Hasil penelitiannya menunjukkan bahwa sebagian besar soal memiliki validitas tinggi dan reliabilitas baik, sehingga instrumen dapat digunakan untuk mengukur keterampilan berpikir kritis siswa. Kajian serupa juga dilakukan oleh Ritdamaya [21] yang mengonstruksi instrumen tes keterampilan berpikir kritis terkait materi suhu dan kalor di tingkat SMA. Instrumen yang dikembangkan terdiri atas 22 soal essay dengan indikator berpikir kritis menurut Ennis, dan hasil analisis menunjukkan validitas serta reliabilitas yang memadai.

Selain itu, penelitian oleh Mardadillah [22] menganalisis butir soal instrumen kemampuan berpikir kritis siswa pada tema usaha kuliner dengan penerapan materi suhu dan kalor. Hasilnya memperlihatkan bahwa instrumen memiliki validitas dan reliabilitas yang baik, serta mampu mengukur keterampilan berpikir kritis siswa secara efektif. Dengan demikian, hasil penelitian ini konsisten dengan kajian terdahulu yang menegaskan bahwa instrumen tes essay berbasis kemampuan berpikir kritis, khususnya pada materi suhu dan kalor, dapat dikembangkan dengan validitas tinggi, reliabilitas baik, daya beda memadai, dan tingkat kesukaran yang proporsional. Hal ini memperkuat keyakinan bahwa instrumen yang dikembangkan layak digunakan sebagai alat evaluasi pembelajaran IPA di SMP, sekaligus memberikan kontribusi terhadap pengembangan asesmen berbasis keterampilan berpikir kritis dalam pendidikan sains.

4. Kesimpulan

Kesimpulan dari penelitian pengembangan instrumen tes essay berbasis kemampuan berpikir kritis pada materi suhu dan kalor dalam pembelajaran IPA di Sekolah Menengah Pertama adalah bahwa instrumen yang terdiri atas 10 butir soal telah memenuhi kriteria kualitas tes yang baik. Hasil analisis menggunakan ANATES menunjukkan bahwa sebagian besar soal memiliki validitas tinggi hingga sangat tinggi, sehingga mampu mengukur keterampilan berpikir kritis siswa sesuai dengan konstruk yang diharapkan. Reliabilitas instrumen berada pada kategori tinggi, menandakan konsistensi internal yang baik dan dapat diandalkan untuk digunakan dalam evaluasi pembelajaran. Analisis daya beda memperlihatkan bahwa mayoritas soal mampu membedakan siswa dengan kemampuan tinggi dan rendah, meskipun terdapat beberapa butir yang perlu revisi agar lebih optimal. Tingkat kesukaran soal terdistribusi secara proporsional dengan dominasi kategori sedang, sehingga tes tidak terlalu mudah maupun terlalu sulit bagi siswa. Secara keseluruhan, instrumen tes essay yang dikembangkan dapat dinyatakan layak digunakan sebagai alat evaluasi pembelajaran IPA di SMP, khususnya untuk mengukur kemampuan berpikir kritis siswa pada materi suhu dan kalor, serta dapat menjadi acuan dalam pengembangan instrumen evaluasi serupa pada materi lain.

Daftar Pustaka

- [1] D. P. Damanik, "Analisis Kemampuan Berpikir Kritis dan Sikap Ilmiah pada Pembelajaran Fisika Menggunakan Model Pembelajaran Inquiry Training (IT) dan Direct Instruction (DI)," PhD Thesis, UNIMED, 2013.
- [2] J. Insiyyah, S. Jumini, dan A. Khoiri, "Implementasi Metode Bahtsul Masail Berbasis Pendidikan Pesantren untuk Meningkatkan Kemampuan Berpikir Kritis dan Kemampuan Menganalisis Peserta Didik pada Pembelajaran Fisika di SMA," *Radiasi J. Berk. Pendidik. Fis.*, vol. 13, no. 2, hlm. 50-54, 2020.
- [3] R. D. Puspitasari, "Integrasi Literasi Digital Dalam Pembelajaran Ipa Fisika Untuk Meningkatkan Kemampuan Berpikir Kritis Siswa," *Unisan J.*, vol. 3, no. 11, hlm. 91-99, 2024.
- [4] D. S. Asysyifa, I. WILUJENG, dan H. Kuswanto, "Analysis of students critical thinking skills using partial credit models (Pcm) in physics learning," *Int. J. Educ. Res. Rev.*, vol. 4, no. 2, hlm. 245-253, 2019.
- [5] U. Sumarmo, W. Hidayat, R. Zukarnaen, H. Hamidah, dan R. Sariningsih, "Kemampuan dan disposisi berpikir logis, kritis, dan kreatif matematik," *J. Pengajaran Mat. Dan Ilmu Pengetah. Alam*, vol. 17, no. 1, hlm. 17-33, 2012.
- [6] N. Akben, "Effects of the problem-posing approach on students' problem solving skills and metacognitive awareness in science education," *Res. Sci. Educ.*, vol. 50, no. 3, hlm. 1143-1165, 2020.
- [7] A. R. Anisa, A. A. Ipungkarti, dan K. N. Saffanah, "Pengaruh kurangnya literasi serta kemampuan dalam berpikir kritis yang masih rendah dalam pendidikan di Indonesia," dipresentasikan pada Current research in education: conference series journal, 2021, hlm. 1-12.
- [8] N. R. Jiwandono, "Kemampuan berpikir kritis (critical thinking) mahasiswa semester 4 (empat) pada mata kuliah psikolinguistik," *Ed-Humanist. J. Ilmu Pendidik.*, vol. 4, no. 1, 2019.

- [9] W. Nadar, "Meningkatkan Kemampuan Berpikir Kritis dengan Penggunaan Strategi Inkuiri: Penelitian Tindakan di TKk Labschool STAI Bani Saleh Bekasi," *J. Ilmu Pendidik. JIP STKIP Kusuma Negara*, vol. 9, no. 2, hlm. 129-144, 2018.
- [10] L. R. Aiken, *Psychological testing and assessment*. Pearson Education India, 2009.
- [11] N. Nurhayati, D. F. Saputri, dan S. L. H. Assegaf, "Pengembangan Instrumen Tes Keterampilan Proses Sains pada Materi Fisika untuk Siswa Sekolah Menengah Pertama," *Edukasi J. Pendidik.*, vol. 17, no. 2, hlm. 145-158, 2019.
- [12] S. Putrawangsa dan U. Hasanah, "Analisis capaian siswa Indonesia pada PISA dan urgensi kurikulum berorientasi literasi dan numerasi," *EDUPEDIKA J. Studi Pendidik. Dan Pembelajaran*, vol. 1, no. 1, hlm. 1-12, 2022.
- [13] I. Kurniati, I. Dewi, dan H. Hasratuddin, "The Development of Student Worksheet Based on PISA to Improve Problem Solving Ability," *Am. J. Educ. Res.*, vol. 6, no. 11, hlm. 1581-1585, Des 2018, doi: 10.12691/education-6-11-18.
- [14] E. S. Kurniawan, "Strategi Jembatan Konsep Analogi Untuk Meningkatkan PhyHOTS Peserta Didik Pada Pembelajaran Fisika," *J. Inov. Pendidik. Sains JIPS*, vol. 4, no. 1, hlm. 26-35, 2023, doi: <https://doi.org/10.37729/jips.v4i1.3023>.
- [15] E. Istiyono, D. Mardapi, dan S. Suparno, "Pengembangan tes kemampuan berpikir tingkat tinggi fisika (pysthots) peserta didik SMA," *J. Penelit. Dan Eval. Pendidik.*, vol. 18, no. 1, hlm. 1-12, 2014.
- [16] R. Jaenudin, U. Chotimah, F. Farida, dan S. Syarifuddin, "Student Development Zone: Higher Order Thinking Skills (Hots) in Critical Thinking Orientation," *Int. J. Multicult. Multireligious Underst.*, vol. 7, no. 9, hlm. 11-19, 2020.
- [17] M. D. Kusuma, U. Rosidin, A. Abdurrahman, dan A. Suyatna, "The Development of Higher Order Thinking Skill (Hots) Instrument Assessment In Physics Study," *IOSR J. Res. Method Educ. IOSRJRME*, vol. 07, no. 01, hlm. 26-32, Jan 2017, doi: 10.9790/7388-0701052632.
- [18] P. Sugiyono, "Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif, dan R&D)(2015 ed.)," *Penerbit Alf. Bdg.*, 2015.
- [19] S. S. Silaban, A. Suhandi, dan Y. E. Gunanto, "Aplikasi Media Simulasi Virtual pada Model Pembelajaran ECIRR untuk Meremediasi Miskonsepsi Siswa pada Materi Perubahan Wujud Zat," dipresentasikan pada Prosiding SNFA (Seminar Nasional Fisika dan Aplikasinya), 2017, hlm. 201-213.
- [20] D. Lestari, L. Berlian, dan L. Berlian, "Pengembangan Instrumen Tes Kemampuan Berpikir Kritis Siswa Kelas VIII Pada Tema Maknanku Kesehatanku:(Development of Critical Thinking Ability Test Instruments for Class VIII Students on the Theme of My Food My Health)," *BIODIK*, vol. 8, no. 2, hlm. 81-89, 2022.
- [21] D. Ritdamaya dan A. Suhandi, "Konstruksi instrumen tes keterampilan berpikir kritis terkait materi suhu dan kalor," *J. Penelit. Pengemb. Pendidik. Fis.*, vol. 2, no. 2, hlm. 87-96, 2016.
- [22] M. Mardatillah, M. C. Sutarja, M. Ahied, W. P. Hadi, dan Y. Yamin, "KEMAMPUAN BERPIKIR KRITIS SISWA MENGGUNAKAN INSTRUMEN SOAL BERBASIS KEARIFAN LOKAL KUE KLEPON PADA MATERI KALOR," *Nat. Sci. Educ. Res. NSER*, vol. 8, no. 1, hlm. 102-110, 2025.