



Jurnal Review Pendidikan dan Pengajaran
<http://journal.universitaspahlawan.ac.id/index.php/jrpp>
 Volume 8 Nomor 1, 2025
 P-2655-710X e-ISSN 2655-6022

Submitted : 01/01/2025
 Reviewed : 01/01/2025
 Accepted : 01/01/2025
 Published : 03/01/2025

Ali Azhar
 Herdiansyah¹
 Dwiky Nurfauzi²
 Nur Assyifa
 Khoirunnisa³
 Hafiziani Eka Putri⁴
 Teten Ginanjar
 Rahayu⁵

PENGEMBANGAN INSTRUMEN UNTUK MENGUKUR KEMAMPUAN BERPIKIR KRITIS SISWA MELALUI KONSEP OPERASI HITUNG PERKALIAN DAN PEMBAGIAN PECAHAN PADA SISWA KELAS VI SD 3 BABAKAN SARI

Abstrak

Penelitian ini bertujuan untuk mengembangkan instrumen evaluasi berpikir kritis siswa kelas 6 SD melalui operasi hitung pembagian. Instrumen ini dirancang dengan konteks kehidupan sehari-hari yang relevan untuk mengevaluasi pemahaman siswa secara komprehensif. Metode penelitian melibatkan uji coba soal pada 15 siswa dengan analisis menggunakan perangkat ANATES untuk menilai validitas, reliabilitas, konteks soal, tingkat kesukaran, dan daya pembeda. Hasil penelitian menunjukkan bahwa instrumen memiliki validitas dan reliabilitas tinggi, tetapi soal sebagian besar berada pada tingkat kesulitan "mudah," sehingga kurang optimal untuk mengukur keterampilan berpikir kritis secara mendalam. Temuan ini menegaskan pentingnya pengembangan soal dengan tingkat kesulitan bertahap, konteks lebih kompleks, dan melibatkan penalaran multistep untuk meningkatkan daya ukur instrumen.

Kata Kunci: Berpikir Kritis, Pembagian Pecahan, Perkalian Pecahan. Evaluasi Pendidikan

Abstract

This study aims to develop an evaluation instrument for critical thinking skills in sixth-grade elementary school students through division operations. The instrument was designed with relevant real-life contexts to comprehensively assess students' understanding. The research method involved testing the questions on 15 students, with data analyzed using ANATES software to evaluate validity, reliability, difficulty levels, level of difficulty, and discrimination indices. The results indicated that the instrument has high validity and reliability; however, most questions were at an "easy" difficulty level, making them less optimal for measuring in-depth critical thinking skills. These findings highlight the importance of developing questions with progressive difficulty levels, more complex contexts, and multi-step reasoning to enhance the instrument's measurement capabilities.

Keywords: Critical Thinking, Division of Fractions, Multiplying Fractions Educational Assesmen

PENDAHULUAN

Kemampuan berpikir kritis merupakan salah satu keterampilan abad ke-21 yang sangat penting untuk dikuasai oleh siswa. Kemampuan ini dapat membantu siswa dalam memahami konsep-konsep akademik dan dalam menghadapi tantangan kehidupan sehari-hari yang semakin kompleks (Simanjuntak, 2019). Di tingkat Sekolah Dasar (SD), kemampuan berpikir kritis sering kali dianggap sebagai dasar yang penting dalam membangun pola pikir analitis siswa.

Pada era modern ini, instrumen evaluasi pendidikan berfungsi sebagai alat untuk mengukur keberhasilan belajar siswa dan sebagai sarana untuk melatih keterampilan berpikir kritis mereka. Salah satu aspek yang menarik untuk diteliti adalah pengembangan instrumen

^{1,2,3,4,5)} Program Studi Pendidikan Guru Sekolah Dasar, Universitas Pendidikan Indonesia kampus Purwakarta

email: ali.azhar393@upi.edu¹, dwiky.fauzi04@upi.edu², nurassyifa.13@upi.edu³, hafizianiekaputri@upi.edu⁴, tetenginanjarr@upi.edu⁵

yang dirancang untuk melatih kemampuan berpikir kritis melalui materi matematika, khususnya operasi hitung pembagian. Matematika sebagai disiplin ilmu dasar memegang peranan penting dalam melatih siswa untuk berpikir logis, sistematis, dan kritis. Namun, dalam praktiknya, pembelajaran matematika sering kali masih berfokus pada aspek penguasaan rumus dan konsep, sehingga aspek pengembangan kemampuan berpikir kritis belum mendapatkan perhatian yang memadai.

Studi ini berfokus pada pengembangan instrumen evaluasi berbasis kemampuan berpikir kritis pada siswa kelas 6 SD di Babakan Sari. Proses ini didukung oleh data yang diolah menggunakan perangkat lunak ANATES dari hasil uji coba soal evaluasi. Data yang terkumpul menunjukkan bahwa tingkat kesulitan soal yang rendah dapat menjadi tantangan dalam mengukur kemampuan berpikir kritis siswa secara optimal.

Berpikir kritis merupakan kemampuan untuk menganalisis informasi secara mendalam, mengevaluasi argumen, serta mengambil keputusan berdasarkan bukti yang ada. Definisi berpikir kritis mencakup kemampuan untuk mengidentifikasi, memahami, dan mengevaluasi informasi dengan pendekatan logis dan objektif. Paul dan Elder (2006) mendefinisikan berpikir kritis sebagai proses disiplin intelektual yang digunakan untuk memahami, mengevaluasi, dan memecahkan masalah. Dalam pendidikan dasar, berpikir kritis penting karena membantu siswa untuk mengembangkan kemampuan analisis, refleksi, dan pengambilan keputusan yang diperlukan untuk pembelajaran di jenjang lebih lanjut. Komponen utama berpikir kritis meliputi interpretasi, analisis, evaluasi, inferensi, eksplanasi, dan regulasi diri. Proses ini melibatkan kemampuan kognitif dan disposisi afektif seperti rasa ingin tahu, berpikir terbuka, dan kesediaan untuk mempertimbangkan sudut pandang lain. Di tingkat pendidikan dasar, keterampilan ini dapat dilatih melalui pendekatan pembelajaran yang mendorong eksplorasi, diskusi, dan penyelesaian masalah nyata (Fadillah, 2024).

Instrumen evaluasi adalah alat yang digunakan untuk mengukur pencapaian belajar siswa, termasuk kemampuan berpikir kritis. Instrumen ini dirancang untuk mengumpulkan data yang valid dan reliabel tentang kemampuan siswa dalam menganalisis, mengevaluasi, dan menyelesaikan masalah. Dalam berpikir kritis, instrumen evaluasi sering kali berbentuk tes yang dirancang secara khusus untuk mengukur aspek-aspek seperti kemampuan inferensi, analisis argumen, dan pengambilan keputusan berbasis data. Agar efektif, instrumen evaluasi harus memenuhi kriteria validitas dan reliabilitas. Validitas memastikan bahwa instrumen benar-benar mengukur kemampuan yang dimaksud, sementara reliabilitas menunjukkan konsistensi hasil pengukuran. Proses pengembangan instrumen melibatkan analisis mendalam tentang kebutuhan evaluasi, perancangan soal yang sesuai, dan uji coba untuk memastikan kualitas instrumen. Dalam pengembangan instrumen berpikir kritis, soal-soal yang digunakan harus mampu merangsang siswa untuk berpikir logis, mempertanyakan asumsi, dan mencari solusi yang berbasis bukti.

Matematika merupakan salah satu disiplin ilmu yang memiliki potensi besar dalam melatih kemampuan berpikir kritis siswa. Sebagai ilmu yang menekankan pada pola, struktur, dan hubungan logis, matematika memberikan ruang bagi siswa untuk mengembangkan keterampilan analisis, evaluasi, dan penyelesaian masalah. Operasi hitung pembagian menuntut pemahaman konsep dasar dan kemampuan untuk menerapkannya dalam berbagai konteks. Hal ini menjadikan matematika sebagai media yang ideal untuk melatih siswa berpikir secara sistematis dan logis. Menurut Kilpatrick, Swafford, dan Findell (2001), pembelajaran matematika yang efektif mencakup lima dimensi utama: pemahaman konsep, kelancaran prosedural, penalaran adaptif, disposisi produktif, dan kompetensi strategis (Sari & Bakhtiar, 2024). Dimensi penalaran adaptif dan kompetensi strategis sangat erat kaitannya dengan kemampuan berpikir kritis. Dengan mengintegrasikan aspek-aspek ini dalam pembelajaran, siswa dapat didorong untuk mengeksplorasi berbagai cara penyelesaian masalah, mempertanyakan asumsi-asumsi yang ada, dan mengembangkan solusi yang kreatif serta berbasis bukti.

ANATES adalah perangkat lunak yang dirancang untuk membantu analisis data dalam pengembangan instrumen evaluasi pendidikan. Perangkat ini memungkinkan pengguna untuk mengukur validitas, reliabilitas, tingkat kesulitan, dan daya pembeda butir soal secara akurat. Dalam pengembangan instrumen berpikir kritis, ANATES sangat berguna untuk memastikan

bahwa soal yang dirancang mampu mengevaluasi kemampuan siswa secara komprehensif. Dengan memanfaatkan hasil analisis ANATES, pengembang instrumen dapat mengidentifikasi butir soal yang perlu direvisi atau disempurnakan (Elviana, 2019). Proses analisis dengan ANATES melibatkan beberapa langkah penting, termasuk input data, pengolahan statistik, dan interpretasi hasil. Perangkat ini menggunakan pendekatan kuantitatif untuk memberikan informasi mendalam tentang karakteristik setiap butir soal, seperti tingkat kesukaran dan korelasi antar soal. Dalam konteks penelitian ini, ANATES digunakan untuk mengevaluasi hasil uji coba soal evaluasi matematika pada siswa kelas 6 SD di Babakan Sari. Hasil analisis menunjukkan bahwa sebagian besar soal memiliki tingkat kesukaran yang tergolong mudah, tetapi tetap signifikan dalam mengukur kemampuan berpikir kritis siswa.

METODE

Penelitian ini dilakukan menggunakan metode kualitatif dengan pendekatan deskriptif untuk menggali secara mendalam proses pengembangan instrumen evaluasi berpikir kritis pada siswa kelas 6 SD. Pendekatan ini bertujuan untuk memahami fenomena dengan memusatkan perhatian pada pengalaman subjek dan konteks pembelajaran. Proses penelitian melibatkan pengumpulan data melalui observasi langsung terhadap proses pembelajaran serta analisis dokumen terkait rancangan soal. Data kualitatif yang terkumpul dianalisis secara tematik untuk mengidentifikasi pola, tema, dan hubungan yang relevan dengan pengembangan instrumen.

Pelaksanaan penelitian dimulai dengan merancang kerangka soal dan pedoman observasi yang disesuaikan dengan tujuan penelitian. Observasi dilakukan selama kegiatan pembelajaran untuk mengamati respon siswa terhadap soal yang diberikan, termasuk kemampuan mereka dalam menganalisis, mengevaluasi, dan menyelesaikan masalah. Seluruh data direkam dan ditranskripsi untuk dianalisis lebih lanjut.

Analisis data dilakukan secara iteratif dengan tahapan reduksi data, penyajian data, dan penarikan kesimpulan. Peneliti mengelompokkan data berdasarkan tema yang berkaitan dengan validitas, relevansi konteks, serta efektivitas soal dalam mengukur kemampuan berpikir kritis. Proses triangulasi dilakukan untuk meningkatkan kredibilitas hasil penelitian, yaitu dengan membandingkan data dari wawancara, observasi, dan dokumen. Hasil analisis ini digunakan untuk mengidentifikasi kekuatan dan kelemahan instrumen yang dirancang, serta memberikan rekomendasi pengembangan lebih lanjut guna meningkatkan kualitas evaluasi berpikir kritis siswa.

HASIL DAN PEMBAHASAN

Soal evaluasi yang digunakan dalam penelitian ini dirancang untuk mengukur kemampuan berpikir kritis siswa kelas 6 SD pada materi operasi hitung pembagian dan perkalian pecahan. Soal-soal tersebut mencakup berbagai konteks kehidupan sehari-hari yang bertujuan membuat siswa berpikir lebih mendalam dalam menerapkan konsep matematika. Berdasarkan analisis, kualitas soal ditinjau dari aspek tingkat kesukaran, daya pembeda, validitas, dan konteks relevansi terhadap kemampuan berpikir kritis.

1. Validitas Butir Soal

Validitas soal dievaluasi dengan menggunakan korelasi antar butir soal. Data ANATES menunjukkan bahwa semua soal memiliki korelasi signifikan dengan skor total, dengan nilai antara 0,718 hingga 0,830.

Hal ini dapat dilihat pada tabel berikut:

Tabel 1. Validitas Butir Soal

No. Soal	Korelasi Butir (r)	Signifikansi
1	0,718	Sangat Signifikan
2	0,748	Sangat Signifikan
3	0,751	Sangat Signifikan
4	0,752	Sangat Signifikan
5	0,830	Sangat Signifikan

Nilai korelasi ini menunjukkan bahwa setiap butir soal memiliki hubungan yang kuat dengan keseluruhan tes. Validitas soal yang tinggi mengindikasikan bahwa semua soal relevan untuk mengukur kemampuan berpikir kritis siswa. Namun, untuk meningkatkan kualitas, soal-soal ini dapat dirancang agar tidak hanya relevan tetapi juga mencakup variasi keterampilan berpikir kritis, seperti evaluasi dan pengambilan keputusan.

Validitas butir soal adalah indikator sejauh mana sebuah soal relevan dengan tujuan tes secara keseluruhan. Dalam analisis yang dilakukan, semua soal memiliki korelasi signifikan terhadap skor total tes, dengan nilai korelasi berkisar antara 0,718 hingga 0,830.

Korelasi ini menunjukkan bahwa setiap soal memiliki hubungan yang kuat dengan tujuan pengukuran, yakni kemampuan berpikir kritis siswa. Nilai korelasi yang tinggi menunjukkan bahwa butir soal dirancang dengan baik untuk mengukur aspek-aspek yang relevan dalam konteks kemampuan siswa.

Meski validitas soal tinggi, penting untuk mencermati elemen yang diukur dalam soal. Sebagian besar soal dalam tes ini cenderung mengukur aspek kemampuan dasar, seperti penerapan konsep pembagian, tanpa mencakup elemen yang lebih kompleks seperti evaluasi keputusan atau interpretasi data. Dalam pengembangan kemampuan berpikir kritis, validitas soal akan lebih optimal jika setiap soal mencakup elemen berpikir tingkat tinggi, seperti mengidentifikasi pola, mengevaluasi solusi alternatif, atau menyintesis informasi dari berbagai sumber.

Untuk meningkatkan validitas instrumen, pengembangan soal di masa depan harus berfokus pada desain soal yang lebih kontekstual dan multiaspek. Soal yang meminta siswa untuk menganalisis data tentang distribusi barang dalam sebuah perusahaan kecil atau soal yang meminta siswa mengevaluasi efisiensi strategi pembagian sumber daya di kelas. Dengan menambahkan dimensi-dimensi ini, soal tidak hanya relevan secara teoritis, tetapi juga mampu mengukur kemampuan berpikir kritis yang lebih mendalam.

2. Reliabilitas Tes

Reliabilitas tes adalah ukuran konsistensi instrumen dalam mengevaluasi kemampuan siswa. Berdasarkan hasil analisis, reliabilitas tes ini mencapai angka 0,77, yang termasuk dalam kategori "cukup tinggi." Hal ini menunjukkan bahwa instrumen memiliki konsistensi yang baik dalam mengukur kemampuan siswa pada materi operasi hitung pembagian. Dengan reliabilitas yang cukup tinggi, hasil tes dapat diandalkan untuk memberikan gambaran yang akurat tentang kemampuan siswa.

Tabel 2. Reliabilitas Tes

Aspek yang Dianalisis	Nilai	Interpretasi
Reliabilitas Tes	0,77	Cukup tinggi
Keterangan	-	Instrumen menunjukkan konsistensi yang baik dalam mengukur kemampuan siswa. Namun, dapat ditingkatkan dengan menambah variasi soal dan tingkat kesulitan.
Jumlah Soal	5	Soal yang terbatas dapat mempengaruhi reliabilitas, karena semakin banyak soal yang digunakan, semakin tinggi konsistensi hasil tes.
Simpang Baku Skor Siswa	3,79	Menunjukkan variasi skor yang moderat antara siswa, yang menunjukkan adanya perbedaan kemampuan yang terukur.

Namun, meski reliabilitas tergolong tinggi, ada beberapa faktor yang dapat memengaruhi konsistensi hasil tes. Salah satunya adalah homogenitas soal yang cenderung berada di tingkat kesulitan "mudah." Soal-soal yang terlalu seragam dapat membuat instrumen kurang sensitif terhadap perbedaan kemampuan siswa, sehingga berdampak pada konsistensi hasil pengukuran. Jumlah soal yang terbatas (hanya lima soal) juga dapat memengaruhi reliabilitas, karena semakin banyak soal yang digunakan, semakin tinggi peluang untuk mendapatkan hasil yang konsisten.

Untuk meningkatkan reliabilitas di masa depan, jumlah soal dalam instrumen perlu ditambah, dengan distribusi tingkat kesulitan yang lebih beragam. Soal dengan konteks yang

lebih menantang dan membutuhkan penalaran multilangkah dapat memperkuat konsistensi pengukuran. Reliabilitas dapat ditingkatkan dengan melakukan uji coba instrumen pada populasi siswa yang lebih besar dan heterogen, sehingga hasil yang diperoleh lebih representatif dan dapat diandalkan.

3. Konteks Soal

Soal evaluasi yang digunakan dalam penelitian ini telah dirancang dengan konteks kehidupan nyata yang menarik dan relevan bagi siswa. Konteks soal mencakup skenario-skenario seperti pembagian luas kebun antara tanaman apel dan jeruk, kebutuhan botol untuk menampung air, serta pengemasan loyang kue ke dalam kotak. Situasi ini dipilih untuk memastikan bahwa soal mudah dipahami dan memiliki keterkaitan langsung dengan pengalaman sehari-hari siswa. Dengan pendekatan ini, siswa diharapkan dapat merasa terhubung dengan soal dan lebih termotivasi untuk menyelesaikannya, karena mereka dapat melihat aplikasi nyata dari konsep matematika yang dipelajari di kelas.

Meskipun relevan, soal-soal ini masih cenderung fokus pada perhitungan langsung tanpa menuntut siswa untuk menerapkan analisis mendalam atau penalaran yang lebih kompleks. Sebagai contoh, soal pembagian luas kebun hanya meminta siswa untuk melakukan perhitungan sederhana, tanpa melibatkan langkah-langkah evaluasi atau perencanaan. Hal ini menunjukkan bahwa instrumen masih lebih mengarah pada pengukuran kemampuan aritmatika dasar daripada keterampilan berpikir kritis yang lebih tinggi. Akibatnya, soal ini kurang mampu melatih siswa untuk menganalisis situasi, membandingkan alternatif, atau membuat keputusan berdasarkan data yang diberikan.

4. Tingkat Kesukaran

Tingkat kesukaran soal menjadi salah satu indikator penting dalam menentukan efektivitas soal sebagai instrumen evaluasi. Data yang diolah melalui ANATES menunjukkan bahwa semua soal berada dalam kategori "mudah". Berikut adalah hasil tingkat kesukaran dari lima soal:

Tabel 3. Tingkat Kesukaran Soal

No. Soal	Rata-Rata Nilai (T)	Tingkat Kesukaran
1	3,70	Mudah
2	7,00	Mudah
3	3,66	Mudah
4	7,00	Mudah
5	4,70	Mudah

Rata-rata tingkat kesukaran yang tergolong rendah menunjukkan bahwa sebagian besar siswa mampu menjawab soal dengan benar. Meskipun ini mengindikasikan pemahaman yang baik terhadap materi, tingkat kesukaran yang seragam menjadi kelemahan karena soal tidak cukup menantang siswa untuk berpikir lebih kritis. Untuk pengembangan ke depan, perlu ditambahkan soal dengan tingkat kesulitan sedang dan sulit guna mengakomodasi variasi kemampuan siswa.

Tingkat kesukaran menjadi salah satu komponen utama dalam menilai kualitas soal. Berdasarkan data yang dianalisis, semua soal yang digunakan dalam evaluasi berada pada kategori "mudah." Hal ini tercermin dari rata-rata skor siswa untuk setiap soal yang berkisar antara 3,66 hingga 7,00.

Secara umum, soal yang berada di kategori "mudah" berarti mayoritas siswa mampu menyelesaikannya dengan baik. Kondisi ini menggambarkan bahwa soal-soal tersebut belum cukup menantang bagi siswa dan cenderung hanya mengukur kemampuan dasar mereka. Tingkat kesukaran seperti ini menjadi perhatian karena tidak sepenuhnya merepresentasikan kemampuan berpikir kritis siswa.

Tingkat kesukaran yang seragam di kategori "mudah" menunjukkan bahwa soal-soal ini kurang mampu mencakup spektrum kemampuan siswa yang lebih luas. Soal yang mudah biasanya hanya mengukur kemampuan siswa untuk mengingat atau memahami informasi, tanpa mengharuskan mereka berpikir lebih mendalam.

Dalam pengembangan kemampuan berpikir kritis, soal-soal dengan kategori "sedang" dan "sulit" diperlukan agar siswa dapat dilatih untuk menganalisis, mengevaluasi, dan menyintesis informasi dalam pemecahan masalah. Sebagai contoh, siswa seharusnya diberikan soal yang memerlukan lebih dari sekadar perhitungan langsung, tetapi juga analisis strategi terbaik dalam situasi tertentu.

Untuk pengembangan instrumen di masa depan, perlu dirancang soal-soal dengan tingkat kesukaran yang lebih beragam. Misalnya, soal dengan tingkat kesulitan "sedang" dapat mencakup kasus-kasus di mana siswa harus membandingkan hasil perhitungan dari beberapa alternatif sebelum menentukan jawaban yang benar. Soal dengan tingkat kesulitan "sulit" dapat menantang siswa untuk membuat keputusan berdasarkan interpretasi data atau konteks tertentu. Dengan adanya distribusi soal yang lebih beragam, kemampuan siswa pada berbagai level kemampuan dapat terukur dengan lebih baik.

5. Daya Pembeda

Daya pembeda adalah aspek lain yang mengukur sejauh mana soal dapat membedakan siswa yang memiliki kemampuan tinggi dan rendah.

Berdasarkan data, daya pembeda kelima soal berkisar antara 35%-45%, dengan rincian sebagai berikut:

Tabel 4. Soal Daya Pembeda

No. Soal	Daya Pembeda (DP%)	Kategori
1	40%	Baik
2	35%	Baik
3	35%	Baik
4	35%	Baik
5	45%	Baik

Semua soal berada dalam kategori "baik," namun nilai daya pembeda soal 2, 3, dan 4 mendekati batas bawah. Hal ini menunjukkan bahwa soal ini masih kurang optimal dalam membedakan kemampuan siswa dengan lebih tajam. Soal dengan daya pembeda yang lebih tinggi perlu dikembangkan, misalnya dengan menambahkan elemen analisis lebih kompleks yang hanya dapat diselesaikan siswa dengan kemampuan berpikir kritis yang lebih baik.

Daya pembeda mencerminkan sejauh mana sebuah soal dapat membedakan siswa yang memiliki kemampuan tinggi dan rendah. Berdasarkan analisis statistik, daya pembeda dari semua soal berada di kategori "baik," dengan nilai berkisar antara 35% hingga 45%.

Soal dengan daya pembeda pada kategori ini sudah cukup mampu membedakan siswa, meskipun nilai pada beberapa soal (khususnya soal 2, 3, dan 4) mendekati batas bawah kategori "baik." Hal ini menunjukkan bahwa instrumen masih memiliki ruang untuk perbaikan dalam hal diferensiasi antara siswa dengan kemampuan berpikir kritis yang berbeda.

Soal dengan daya pembeda rendah dapat disebabkan oleh desain soal yang terlalu sederhana atau jawaban yang terlalu mudah ditebak. Soal yang hanya menuntut perhitungan matematis sederhana mungkin tidak cukup menantang siswa dengan kemampuan berpikir kritis yang lebih tinggi. Hal ini menyebabkan perbedaan antara siswa dengan kemampuan tinggi dan rendah menjadi tidak terlalu terlihat. Jika distraktor (pilihan jawaban salah) tidak dirancang dengan baik, siswa dengan kemampuan rendah pun dapat dengan mudah menemukan jawaban yang benar, sehingga daya pembeda soal menurun.

Untuk meningkatkan daya pembeda, instrumen harus dirancang dengan memasukkan elemen yang memerlukan analisis dan sintesis lebih kompleks. Soal dengan konteks realistik dan multistep dapat menjadi solusi. Soal yang mengharuskan siswa menentukan strategi terbaik dalam pembagian sumber daya atau menghitung pembagian yang optimal dalam suatu skenario akan menuntut kemampuan berpikir tingkat tinggi.

Temuan Penting

Salah satu temuan utama dari analisis statistik soal adalah keseragaman tingkat kesukaran. Berdasarkan hasil evaluasi, semua soal tergolong dalam kategori "mudah," yang menunjukkan bahwa sebagian besar siswa mampu menjawabnya dengan benar tanpa banyak kesulitan. Meski ini menandakan bahwa materi pembelajaran sudah dikuasai dengan baik oleh

siswa, instrumen seperti ini tidak memberikan tantangan yang cukup untuk mengukur kemampuan siswa. Tingkat kesukaran yang seragam juga membuat instrumen kurang mampu membedakan siswa dengan tingkat kemampuan yang beragam. Pengembangan instrumen di masa depan perlu mencakup variasi tingkat kesulitan, yaitu soal-soal dengan kategori "sedang" dan "sulit."

Konteks soal menjadi poin penting dalam temuan ini. Meskipun soal-soal sudah dirancang dengan memasukkan elemen kontekstual, seperti perhitungan dalam kehidupan sehari-hari, mereka sebagian besar masih berfokus pada perhitungan langsung. Soal seperti ini kurang menantang siswa untuk menggunakan kemampuan berpikir kritis yang lebih mendalam, seperti analisis atau evaluasi informasi. Dalam pembelajaran abad ke-21, penting untuk mengintegrasikan elemen kehidupan nyata yang memerlukan pemecahan masalah dan pengambilan Keputusan (Tarihoran, 2019). Sebagai contoh, soal dapat dirancang untuk meminta siswa merencanakan distribusi sumber daya dalam konteks tertentu atau mengevaluasi strategi pembagian terbaik berdasarkan beberapa variabel.

Peningkatan daya pembeda juga menjadi perhatian dalam analisis ini, khususnya pada soal tertentu seperti soal nomor 2 dan 4. Daya pembeda yang mendekati batas bawah kategori "baik" menunjukkan bahwa soal-soal ini kurang optimal dalam membedakan siswa dengan kemampuan tinggi dan rendah. Soal-soal ini perlu dirancang ulang dengan menambahkan elemen kompleksitas atau distraktor yang lebih efektif untuk menguji kemampuan siswa secara lebih mendalam. Sebagai contoh, soal nomor 2 dapat dimodifikasi menjadi soal multi-langkah yang melibatkan analisis atau pengambilan keputusan, sementara soal nomor 4 dapat mencakup skenario yang memerlukan interpretasi data sebelum melakukan perhitungan. Perubahan ini diharapkan dapat meningkatkan daya pembeda sekaligus memberikan gambaran lebih akurat tentang tingkat kemampuan siswa.

Temuan penting lainnya adalah soal-soal yang lebih cenderung mengukur kemampuan aritmatika dasar dibandingkan keterampilan berpikir kritis. Hal ini menjadi tantangan dalam pengembangan instrumen untuk evaluasi kemampuan berpikir kritis. Soal-soal yang hanya menuntut perhitungan matematis sederhana tidak cukup untuk melatih siswa dalam menganalisis, mengevaluasi, atau menyintesis informasi. Pengembangan soal di masa depan perlu memperluas cakupan kemampuan yang diukur. Soal dapat dirancang untuk meminta siswa mengevaluasi efektivitas suatu strategi, membandingkan hasil dari berbagai pendekatan, atau membuat keputusan berdasarkan interpretasi data. Pendekatan ini dapat meningkatkan kualitas soal dan mendukung pengembangan keterampilan berpikir kritis siswa secara komprehensif.

SIMPULAN

Berdasarkan hasil analisis terhadap instrumen evaluasi yang digunakan, dapat disimpulkan bahwa kualitas soal secara umum sudah memenuhi kriteria validitas dan reliabilitas yang baik, tetapi masih memerlukan perbaikan dalam aspek tingkat kesukaran, daya pembeda, dan fokus pengukuran keterampilan berpikir kritis. Seluruh soal yang dianalisis berada dalam kategori "mudah," yang menunjukkan bahwa siswa mampu menyelesaikan soal dengan baik. Namun, keseragaman tingkat kesukaran ini menjadi kelemahan karena tidak memberikan variasi tantangan yang cukup untuk mengukur kemampuan siswa dengan berbagai tingkat pemahaman.

Daya pembeda pada beberapa soal berada di batas bawah kategori "baik," sehingga soal-soal ini perlu dirancang ulang agar lebih efektif dalam membedakan siswa dengan kemampuan tinggi dan rendah. Meski validitas setiap butir soal cukup tinggi, fokus soal masih cenderung mengukur kemampuan aritmatika dasar tanpa menantang siswa untuk menggunakan keterampilan berpikir kritis secara mendalam.

Untuk meningkatkan kualitas instrumen, perlu dilakukan pengembangan soal dengan tingkat kesulitan yang lebih beragam dan konteks yang lebih kompleks. Soal-soal dengan tingkat kesulitan sedang dan sulit dapat dirancang untuk mencakup elemen analisis, evaluasi, dan pengambilan keputusan.

Integrasi konteks kehidupan nyata dalam soal perlu ditingkatkan dengan memasukkan skenario multi langkah yang mendorong siswa untuk berpikir secara logis dan kritis. Instrumen yang lebih beragam dan kontekstual akan memberikan hasil evaluasi yang lebih akurat dan

berkontribusi pada pengembangan keterampilan berpikir kritis siswa, yang merupakan salah satu tujuan utama dalam pembelajaran abad ke-21.

Saran untuk Pengembangan Instrumen Selanjutnya

- **Menambah Variasi Soal**

Instrumen evaluasi perlu dikembangkan dengan menambahkan soal yang memiliki tingkat kesulitan bertahap, mulai dari kategori "mudah," "sedang," hingga "sulit." Soal dengan tingkat kesulitan yang beragam akan memungkinkan pengukuran kemampuan siswa secara lebih menyeluruh, mencakup siswa dengan tingkat pemahaman yang beragam. Setiap kategori kesulitan perlu dirancang untuk melibatkan keterampilan berpikir kritis yang berbeda, seperti membandingkan hasil perhitungan, menganalisis penyebab perbedaan hasil, dan menyimpulkan solusi terbaik dari beberapa alternatif (Magdalena, 2023).

- **Meningkatkan Kompleksitas Konteks Soal**

Pengembangan soal sebaiknya mengintegrasikan permasalahan nyata yang menantang siswa untuk mengambil keputusan berdasarkan analisis pembagian. Siswa dapat diberikan skenario tentang alokasi sumber daya dalam sebuah proyek, seperti menentukan jumlah bahan yang dibutuhkan untuk setiap tahap pembangunan rumah. Soal semacam ini dapat mengasah keterampilan berpikir kritis dan memberikan siswa pemahaman tentang aplikasi praktis dari konsep pembagian dalam kehidupan sehari-hari (Fathoni *et al* 2023).

- **Penggunaan Soal dengan Penalaran Multilangkah**

Soal perlu dirancang untuk melibatkan lebih dari satu langkah pemikiran, sehingga siswa harus melalui proses berjenjang dalam menyelesaikan persoalan. Soal dapat meminta siswa menghitung pembagian awal, kemudian menggunakan hasil tersebut untuk memecahkan masalah berikutnya dalam konteks yang sama. Siswa dilatih untuk berpikir secara sistematis dan mempertimbangkan konsekuensi dari setiap langkah yang mereka ambil (Astari, 2023).

- **Evaluasi dan Revisi Berkelanjutan**

Proses pengembangan instrumen harus melibatkan evaluasi yang berkelanjutan dengan menguji kan soal kepada kelompok siswa yang lebih heterogen, baik dari segi tingkat kemampuan maupun latar belakang. Hasil dari uji coba ini dapat digunakan untuk merevisi instrumen sehingga lebih sesuai dengan kebutuhan pembelajaran dan lebih akurat dalam mengukur kemampuan siswa. Kolaborasi dengan guru dan ahli pendidikan dapat dilakukan untuk memastikan bahwa soal tidak hanya valid secara statistik, tetapi juga relevan secara pedagogis.

Pengembangan instrumen dengan memperhatikan saran-saran ini diharapkan dapat menghasilkan evaluasi yang lebih efektif, komprehensif, dan sesuai dengan kebutuhan pembelajaran modern.

Soal yang di uji cobakan

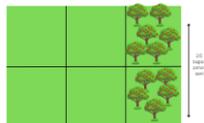
LEMBAR SOAL TES

Nama :

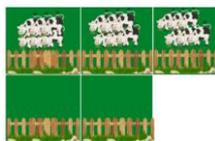
No Absen :

Tanggal :

1. Pak Ahmad mempunyai sebuah kebun memiliki luas $\frac{5}{6}$ hektare. Dari kebun tersebut, $\frac{2}{5}$ di tanami pohon apel, dan sisanya di tanami pohon jeruk. Berapa luas kebun pak Ahmad yang ditanami pohon apel?



2. Sebuah peternakan memiliki luas $\frac{4}{5}$ hektare. Sebagian besar peternakan, yaitu $\frac{3}{4}$, digunakan untuk kandang sapi, sedangkan sisanya digunakan untuk kandang kambing. Hitunglah berapa luas peternakan yang digunakan untuk kandang sapi!



3. Ibu Siti sedang menyiapkan kue untuk acara keluarga besar. Ia memiliki 6 loyang kue yang akan dikemas ke dalam kotak. Setiap kotak dapat menampung $\frac{3}{4}$ loyang kue. Hitunglah jumlah kotak yang dibutuhkan untuk mengemas semua kue ?.



4. Dina sedang mengisi botol-botol air minum untuk piknik bersama teman-temannya. Setiap botol dapat menampung liter $\frac{3}{2}$ air. Dina memiliki 8 botol kosong yang siap diisi. Berapa liter air yang diperlukan untuk mengisi semua botol tersebut jika setiap botol terisi penuh ?.



5. Ani memiliki $\frac{3}{4}$ liter susu. Ia ingin menuangkan susu tersebut ke dalam gelas-gelas kecil, di mana setiap gelas dapat menampung $\frac{1}{8}$ liter susu. Berapa banyak gelas yang dapat Ani isi penuh dengan susu tersebut?



Berikut adalah data siswa beserta skornya:

Tabel 5. Data siswa

No. Urut	Nama Siswa	Skor Ganjil	Skor Genap	Skor Total
1	Agustina Nur Saftiri	8	6	14
2	Ahmad Syaibatul Kamdi R.	15	10	25
3	Akila Safira Putri	15	7	22
4	Aprilia Nur Meylani	13	9	22
5	Azka Fathurizky	12	9	21
6	Daren Leopold Lukmawar	12	7	19
7	Dimas Abyasa	8	7	15
8	Gadis Fitria Nugroho	8	7	15
9	Luthfi Luqman Hakim A.	14	7	21
10	Muhammad Evan Hafizh N.	14	8	22
11	Navika Alike Putri	15	10	25
12	Sabia Fatisha Hilmi	15	10	25
13	Sakya Anantia Putri	11	7	18
14	Vicky Zevar Arvalo Putra	10	7	17
15	Zahra Arsyifa	12	10	22

DAFTAR PUSTAKA

Astari, T. (2023). *Pengembangan Lks Matematika Realistik Di Sekolah Dasar*. Cv. Edupedia Publisher.

ELVIANA, S. A. (2019). Analisis Butir Soal Evaluasi Pembelajaran PAI Menggunakan Program Anates di FTK UIN Ar-Raniry.

Fadillah, N. A. (2024). *PENGARUH MODEL PEMBELAJARAN DISCOVERY LEARNING TERHADAP KETERAMPILAN BERPIKIR KRITIS PESERTA DIDIK PADA MATA PELAJARAN IPAS KELAS IV SD (Penelitian Quasi Eksperimen di SDN 067 Nilem Tahun Ajaran 2022/2023)* (Doctoral dissertation, FKIP UNPAS).

- Fathoni, A., Prasodjo, B., Jhon, W., & Zulqadri, D. M. (2023). *Media dan Pendekatan Pembelajaran di Era Digital: Hakikat, Model Pengembangan & Inovasi Media Pembelajaran Digital*.
- Magdalena, I. (2023). *Evaluasi Pembelajaran Sekolah Dasar*. CV Jejak (Jejak Publisher).
- Sari, A. D. I., & Bakhtiar, A. M. (2024). ANALISIS KEMAMPUAN PEMAHAMAN KONSEP DAN KETERAMPILAN PROSEDURAL MATA PELAJARAN MATEMATIKA KELAS 2 UPT SDN 100 GRESIK. *Trigonometri: Jurnal Matematika dan Ilmu Pengetahuan Alam*, 1(3), 54-64.
- Simanjuntak, M. D. R. (2019). *Membangun Ketrampilan 4 C Siswa Dalam Menghadapi Revolusi Industri 4.0*.
- Tarihoran, E. (2019). Guru dalam pengajaran abad 21. *Sapa: Jurnal Kateketik Dan Pastoral*, 4(1), 46-58.