



## Menganalisis Soal Evaluasi Bilangan Cacah Kelas 4 SDN Jatisari 1 Melalui Hasil Analisis Anates

*Analyzing the Evaluation Questions of Whole Numbers for Class 4 SDN Jatisari 1 Through the Results of Anates Analysis*

Mochammad Fadhlan Dzirkullah<sup>1\*</sup>, Muhammad Anji Rijal<sup>2</sup>, Viona Christina Anatasya Lumbanraja<sup>3</sup>, Teten Ginanjar Rahayu<sup>4</sup>, Hafiziani Eka Putri<sup>5</sup>

<sup>1</sup> PGSD, Kampus UPI Purwakarta, fdlnzirkull.2911@upi.edu

<sup>2</sup> PGSD, Kampus UPI Purwakarta, anji.rijal12@upi.edu

<sup>3</sup> PGSD, Kampus UPI Purwakarta, vionatasya.14@upi.edu

<sup>4</sup> PGSD, Kampus UPI Purwakarta, tetenginjarr@upi.edu

<sup>5</sup> PGSD, Kampus UPI Purwakarta, hafizianiekaputri@upi.edu

\*Corresponding Author: E-mail: [fdlnzirkull.2911@upi.edu](mailto:fdlnzirkull.2911@upi.edu)

### Artikel Penelitian

#### Article History:

Received: 18 Nov, 2024

Revised: 21 Dec, 2024

Accepted: 29 Jan, 2025

#### Kata Kunci:

Bilangan Cacah;

Daya Pembeda;

Efektivitas Pengecoh;

Reliabilitas;

Tingkat Kesukaran;

Validitas

#### Keywords:

*Distinguishing Power;*

*Effectiveness Of Distractors;*

*Level Of Difficulty;*

*Reliability;*

*Validity;*

*Whole Numbers*

### ABSTRAK

Penelitian ini bertujuan untuk menganalisis kualitas butir soal evaluasi pembelajaran matematika pada materi bilangan cacah untuk siswa kelas 4 SD, dengan fokus pada validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Instrumen evaluasi yang berkualitas diperlukan untuk memastikan pengukuran kompetensi siswa secara akurat. Penelitian menggunakan metode deskriptif kuantitatif dengan pengumpulan data melalui observasi, dokumentasi, dan tes. Data diambil dari 10 siswa SDN Jatisari 1, Kecamatan Jatisari Kabupaten Karawang, yang mengerjakan 5 butir soal uraian. Analisis dilakukan menggunakan software Anates versi 4.0 untuk mengevaluasi kualitas butir soal berdasarkan kriteria yang telah ditetapkan. Hasil penelitian menunjukkan bahwa besar butir soal memenuhi kriteria validitas dan reliabilitas, dengan koefisien reliabilitas kategori tinggi ( $\geq 0,7$ ). Tingkat kesukaran soal cenderung bervariasi, memungkinkan distribusi hasil yang seimbang. Daya pembeda sebagian besar soal memadai, sementara pengecoh efektif dalam mengalihkan pilihan siswa. Simpulan dari penelitian ini adalah butir soal yang dianalisis memiliki kualitas yang cukup baik dan dapat digunakan sebagai acuan dalam penyusunan instrumen evaluasi pembelajaran yang lebih efektif di masa mendatang. Temuan ini diharapkan membantu pendidik dalam mengembangkan instrumen evaluasi yang valid, reliabel, dan komprehensif.

### ABSTRAK

*This study aims to analyze the quality of mathematics learning evaluation items on whole number material for grade 4 elementary school students, with a focus on validity, reliability, level of difficulty, distinguishing power and effectiveness of distractors. Quality evaluation instruments are needed to ensure accurate measurement of student competency. The research uses quantitative descriptive methods by collecting data through observation, documentation and tests. Data was taken from 10 students of SDN Jatisari 1, Jatisari District, Karawang Regency, who worked on 5 essay questions. Analysis was carried out using Anates software version 4.0 to evaluate the quality of the questions based on predetermined criteria. The research results showed that the item size met the validity and reliability criteria, with a high category reliability coefficient ( $\geq 0.7$ ). The difficulty level of questions tends to vary, allowing for a balanced distribution of results. The discriminating power of most questions was adequate, while distractors were effective in shifting students' choices. The conclusion of this research is that the items analyzed are of fairly good quality and can be used as a reference in preparing more effective learning evaluation instruments in the future. These findings are expected to help educators in developing valid, reliable and comprehensive evaluation instruments*

DOI: [10.56338/jks.v8i1.6724](https://doi.org/10.56338/jks.v8i1.6724)

## PENDAHULUAN

Penilaian hasil belajar merupakan salah satu aspek penting dalam proses pembelajaran, yang berfungsi untuk mengevaluasi pencapaian kompetensi siswa. Salah satu cara untuk mengetahui tingkat keberhasilan pembelajaran adalah melalui penyusunan dan penggunaan instrumen penilaian yang baik dan berkualitas. Instrumen yang berkualitas harus memenuhi beberapa kriteria, seperti validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Kriteria-kriteria tersebut membantu memastikan bahwa instrumen mampu memberikan gambaran yang akurat mengenai kemampuan siswa. Pada jenjang Sekolah Dasar, khususnya dalam mata pelajaran Matematika, pemahaman terhadap konsep dasar seperti bilangan cacah menjadi fondasi penting bagi perkembangan kemampuan siswa di jenjang berikutnya. Oleh karena itu, diperlukan instrumen evaluasi yang dapat membantu para pendidik dalam menentukan butir soal yang baik dan yang layak untuk dipertahankan, serta butir soal yang harus diperbaiki atau bahkan tidak dapat dipakai (Mustaqim, 2024).

Penelitian ini bertujuan untuk menganalisis kualitas butir soal Evaluasi Penilaian Bilangan Cacah pada siswa kelas 4 SD. Pendekatan yang digunakan adalah deskriptif kuantitatif untuk menganalisis lima butir soal berbentuk uraian yang dikerjakan oleh sepuluh siswa. Data dikumpulkan melalui observasi, dokumentasi, dan tes, lalu dianalisis menggunakan software Anates 4.0 untuk mengukur validitas, reliabilitas, tingkat kesukaran, daya pembeda soal, serta efektivitas pengecoh. Diharapkan, hasil analisis ini dapat memberikan umpan balik yang berguna untuk perbaikan dan penyempurnaan butir soal yang akan digunakan dalam tes di sekolah, seperti yang dijelaskan oleh Mochammad Noor Akhmadi (2021) dalam Jurnal Analisis Butir Soal Pas Matematika Peminatan: Daya Pembeda, Tingkat Kesukaran, dan Kualitas Pengecoh (Mustaqim, 2024).

## METODE

Dalam penelitian ini menggunakan metode deskriptif kuantitatif. Jenis penelitian deskriptif kuantitatif meliputi pengumpulan data menggunakan instrumen soal berbentuk uraian (Aqila, D. S. 2022). Penelitian ini diambil dari siswa kelas 4 SDN Jatisari 1 kecamatan Jatisari kabupaten Karawang. Dalam menghitung analisis butir soal dan hasil jawaban penilaian, data yang digunakan dalam penilaian ini adalah data Evaluasi Penilaian Bilangan Cacah kelas 4 sd sebanyak 5 butir soal. Objek dalam penelitian ini diteliti berdasarkan hasil analisis 5 butir soal dengan jumlah 10 siswa.

Teknik dalam pengumpulan data yaitu observasi, dokumentasi dan tes. Metode yang dapat digunakan dalam melakukan analisis butir soal dilakukan secara kuantitatif. Menurut (Fiska et al., 2021) aspek yang perlu diperhatikan dalam mengetahui kualitas butir soal yang baik, dimulai dari validitas, daya beda, tingkat kesukaran, efektivitas pengecoh dan reliabilitas pada setiap butir soal. Analisis kuantitatif dilakukan untuk mengevaluasi kualitas butir soal, mencakup validitas, reliabilitas, tingkat kesukaran, dan efektivitas pengecoh. Proses analisis menggunakan software Anates, dengan fokus pada pengukuran validitas, reliabilitas, tingkat kesukaran, dan daya pembeda butir soal Evaluasi Penilaian Bilangan Cacah untuk siswa kelas 4 SD.

## HASIL DAN PEMBAHASAN

Penelitian ini menggunakan pendekatan deskriptif kuantitatif. Jenis penelitian deskriptif kuantitatif ini mencakup pengumpulan data dengan menggunakan instrumen soal yang berbentuk uraian. Penelitian ini diambil dari siswa kelas 4 SDN Jatisari 1 kecamatan Jatisari kabupaten Karawang. Dalam menghitung analisis butir soal dan hasil jawaban penilaian, data yang digunakan dalam penilaian

ini adalah data Evaluasi Penilaian Bilangan Cacah kelas 4 SDN sebanyak 5 butir soal dengan tes yang dibuat dalam bentuk uraian.

Teknik pengumpulan data yang digunakan dalam penelitian ini meliputi observasi, dokumentasi, dan tes. Analisis butir soal dilakukan dengan metode kuantitatif, yang mencakup aspek validitas, reliabilitas, tingkat kesukaran, dan efektivitas pengecoh. Hasil yang diperoleh kemudian dianalisis menggunakan software Anates. Analisis kuantitatif ini bertujuan untuk menilai kualitas butir soal, dengan memperhatikan aspek validitas, reliabilitas, tingkat kesukaran, dan daya pembeda pada butir soal Evaluasi Penilaian Bilangan Cacah untuk kelas 4 SD.

Terdapat 5 butir soal yang akan diuji pada software Anates dengan mata pelajaran Matematika pada tema bilangan cacah. Soal-soal ini dirancang untuk menguji kemampuan siswa dalam memahami, menganalisis, mengevaluasi, serta mengaplikasikan konsep yang telah dipelajari. Pada bagian berikut, disajikan butir soal secara rinci, yang disusun berdasarkan hasil kajian dan analisis materi terkait:

Kisi - Kisi:

1. Soal ini menguji pemahaman tentang operasi bilangan bulat, terutama penjumlahan dan pengurangan, untuk menghitung perubahan posisi lift. Indikator yang diharapkan adalah kemampuan siswa menghitung posisi lift setelah beberapa perubahan lantai, baik naik maupun turun. Soal ini melibatkan level kognitif C4 (Analisis) di mana siswa diharapkan dapat mengingat konsep dasar, memahami cara perhitungan, dan menerapkan konsep tersebut dalam soal yang berurutan.
2. Soal ini menguji kemampuan siswa dalam operasi bilangan bulat (penjumlahan dan pengurangan) untuk menghitung jumlah kelereng yang dimiliki Andi setelah beberapa transaksi. Siswa harus menganalisis perubahan jumlah kelereng Andi setelah ia memberi sebagian kepada Budi, menerima dari Cici, dan membagikan sisa kelereng kepada empat temannya secara merata. Indikator yang diuji adalah kemampuan menganalisis perubahan jumlah dan menghitung pembagian secara merata, dengan level kognitif C5 (Mengevaluasi).
3. Soal ini menguji kemampuan siswa dalam menghitung perubahan jumlah kelereng Dina melalui operasi bilangan bulat dan pembagian. Siswa harus menganalisis pengurangan, penambahan, dan pembagian kelereng Dina secara merata kepada teman-temannya, serta mengevaluasi apakah jumlah tersebut cukup untuk dibagi secara adil. Indikator yang diuji adalah kemampuan menghitung dan mengevaluasi pembagian kelereng, dengan level kognitif C5 (Mengevaluasi).
4. Soal ini menguji kemampuan siswa dalam operasi hitung bilangan bulat, seperti penjumlahan, pengurangan, dan pembagian. Siswa harus menghitung total pohon mangga yang dimiliki Anthony setelah pemindahan dan pembelian, kemudian menentukan jumlah yang bisa dibagikan merata kepada delapan petani dan menghitung sisa pohon setelah pembagian. Indikator yang diuji adalah kemampuan menganalisis perubahan jumlah pohon dan menentukan hasil pembagian, dengan level kognitif C6 (Mencipta).
5. Soal ini menguji kemampuan siswa dalam menghitung ketinggian akhir pesawat menggunakan operasi bilangan bulat. Siswa harus menerapkan konsep penjumlahan dan pengurangan untuk menentukan ketinggian setelah perubahan bertahap. Indikator yang diuji adalah kemampuan menerapkan operasi hitung dalam konteks perubahan ketinggian, dengan level kognitif C5 (Mengevaluasi), C6 (Mencipta).

Soal – Soal :

1. Sebuah lift berada di lantai 7. Lift tersebut: Turun 3 lantai Naik 5 lantai Turun lagi 2 lantai Setelah itu, ada seseorang yang masuk, dan lift naik lagi 4 lantai. Tentukan di lantai berapa lift berada setelah setiap perubahan!
2. Andi memiliki 24 butir kelereng. Ia memberikan 8 butir kepada Budi, kemudian mendapatkan 4 butir dari Cici. Setelah itu, Andi membagikan kelerengnya kepada 4 teman secara merata. Berapa jumlah kelereng yang masing-masing teman Andi dapatkan?
3. Dina memiliki 200 butir kelereng. Ia memberikan 50 butir kepada 5 temannya secara merata. Setelah itu, Dina membeli 30 butir kelereng lagi dan membagikan semuanya kepada 6 temannya. Apakah jumlah kelereng yang dimiliki Dina cukup untuk dibagikan merata kepada 6 temannya?
4. Anthony memiliki 1200 pohon mangga yang ditanam dalam 10 barisan. Setelah memindahkan 300 pohon, petani membeli 500 pohon lagi dan menanamnya dalam 5 barisan. Anthony ingin membagikan sisa pohon mangga secara merata kepada 8 petani di desa. hitunglah jumlah pohon mangga yang bisa di bagikan pa Anthony kepada 8 petani di desa dan berapa banyak pohon mangga yang tersisa setelah pembagian.
5. Pesawat Garuda terbang pada ketinggian 8.000 meter di atas permukaan laut. Kemudian pesawat turun sejauh 2.500 meter. Lalu, pesawat naik lagi sejauh 1.200 meter. Berapakah ketinggian pesawat sekarang?

**Validitas**

Validitas adalah konsep dasar yang sangat penting untuk diingat ketika membuat penilaian atau evaluasi apa pun, karena validitas mengacu pada sejauh mana suatu tes, alat ukur, atau metode evaluasi benar-benar mengukur apa yang dimaksudkan untuk diukur, sehingga hasil yang diperoleh dapat dianggap akurat, relevan, dan dapat dipercaya dalam konteks yang dimaksudkan. Uji validitas dilakukan untuk menilai apakah sebuah kuesioner dapat dianggap sah atau valid. Sebuah kuesioner dianggap valid jika pertanyaan-pertanyaan di dalamnya mampu menggambarkan atau mengukur apa yang seharusnya diukur oleh kuesioner tersebut.

**Reliabilitas**

Menurut Ruffi'i (2020) Mengukur serta mengamati objek yang diukur yang dilakukan dengan konsisten pada sebuah tes disebut reliabilitas. Sebuah tes dianggap konsisten jika hasil yang diperoleh hampir sama atau identik ketika tes tersebut diujikan berulang kali. Semakin tinggi koefisien reliabilitas yang (mendekati angka 1)maka semakin besar pula tingkat konsistensi dari tes tersebut. Reliabilitas dinyatakan dapat dipercaya jika hasilnya relatif stabil ketika diuji beberapa kali. Oleh karena itu, jika suatu tes untuk menguji aspek reliabilitas hanya dilakukan sekali, hasilnya belum cukup bisa diandalkan. Reliabilitas yang baik tergolong dalam kategori tinggi, yang menunjukkan tingkat konsistensi 0,7. Menekankan bahwa reliabilitas juga mencakup stabilitas hasil yang dapat digeneralisasikan ke berbagai individu.

Tabel 1. Tabel Reabilitas

Rata <sup>2</sup> 15,50	= Simpangan Baku = 2,17	Kolerasi XY = 0,60	Relialibitas Tes = 0,75	Rata <sup>2</sup> = 15,50	
No.Urut	No.Subyek	Kode>Nama Subyek	Skor Ganjil	Skor Genap	Skor Total
1	7	Fajar	9	7	16

2	2	Muhammad Alf	9	5	14
3	3	Syaqia	8	6	14
4	8	Guhmita	8	6	14
5	9	Syamiah	7	7	14
6	4	Halfah	5	7	12
7	5	Indah R	8	4	12
8	1	Zahra	6	5	11
9	10	Dafa	7	3	10
10	6	Qunsya	5	4	9

### Tingkat Kesukaran

Tabel ini menjelaskan tingkat kesukaran dari setiap butir soal berdasarkan persentase keberhasilan peserta tes dalam menjawab dengan benar. Tingkat kesulitan suatu soal merujuk pada nilai numerik yang menggambarkan seberapa sulit sebuah soal tes (Fietri, 2021). Tingkat kesulitan ini ditentukan dengan membandingkan jumlah total siswa dengan jumlah siswa yang menjawab benar pada setiap soal. Soal dianggap semakin mudah jika nilai indeks kesukarannya mendekati 1,00. Tingkat kesukaran dapat di akolasikan jadi tiga kategori utama, yaitu Sangat Mudah, Mudah, dan Sedang, yang masing-masing memiliki kriteria tertentu. Setiap kategori menggambarkan bagaimana soal tersebut diterima dan dikerjakan oleh peserta tes, mulai dari soal yang hampir semua peserta dapat menjawab dengan benar hingga soal yang menimbulkan lebih banyak tantangan.

Soal dalam kategori sangat mudah memiliki tingkat kesukaran lebih dari 90%, yang berarti lebih dari 90% peserta tes dapat menjawabnya dengan benar. Soal ini biasanya bersifat sederhana, menguji pengetahuan dasar, atau mencakup materi yang sudah dikuasai mayoritas peserta. Dalam tes, soal kategori ini sering digunakan untuk memastikan peserta memahami konsep dasar sebelum melanjutkan ke soal yang lebih sulit. Tingkat keberhasilan yang sangat tinggi ini menunjukkan bahwa soal tersebut tidak memberikan kesulitan berarti bagi peserta. Namun, terlalu banyak soal sangat mudah dalam sebuah tes dapat mengurangi efektivitasnya dalam membedakan kemampuan peserta.

Sementara itu, soal mudah memiliki tingkat kesukaran antara 70% hingga 90%, yang berarti mayoritas peserta mampu menjawabnya dengan benar. Soal-soal ini lebih menantang dibandingkan soal sangat mudah, tetapi tetap sederhana bagi peserta yang memiliki pemahaman cukup baik. Soal dalam kategori mudah sering kali menjadi mayoritas dalam sebuah tes, karena mampu mengevaluasi peserta dengan tingkat kesulitan yang moderat. Soal mudah dirancang untuk mengukur kompetensi yang sedikit lebih kompleks dibandingkan soal sangat mudah.

Adapun soal sedang memiliki tingkat kesukaran antara 50% hingga 70%, di mana hanya sekitar separuh hingga dua pertiga peserta yang dapat menjawab dengan benar. Tingkat keberhasilan yang lebih rendah dibandingkan kategori sebelumnya menunjukkan bahwa soal-soal ini berbeda tingkat kesulitannya dan membutuhkan pemahaman yang lebih mendalam atau kemampuan analisis dari peserta. Soal sedang biasanya digunakan dalam tes untuk mengukur keterampilan peserta dalam menjawab pertanyaan yang lebih kompleks dan menantang. Soal dalam kategori ini sangat penting dalam tes karena membantu membedakan peserta dengan tingkat kemampuan yang beragam. Namun, jika terlalu sedikit soal sedang, tes dapat menjadi terlalu mudah, sedangkan jika terlalu banyak, tes dapat menjadi terlalu sulit bagi sebagian besar peserta (Fajriah et al., 2020; Putri & Warmi, 2022).

Ketiga kategori ini saling melengkapi dalam sebuah tes, dan proporsinya harus disesuaikan dengan tujuan tes untuk menghasilkan penilaian yang valid dan informatif.

Tabel 2. Tingkat Kesukaran

● Jml Subyek = 10		● Butir Soal = 5	
● No Butir Baru	● No Butir Asli	● Tkt. Kesukaran(%)	● Tafsiran
● 1	● 1	● 66.67	● Sedang
● 2	● 2	● 91.67	● Sangat Mudah
● 3	● 3	● 91.67	● Sangat Mudah
● 4	● 4	● 66.67	● Sedang

Daya pembeda adalah kemampuan sebuah butir soal untuk membedakan antara siswa yang memiliki keterampilan tinggi (kelompok atas) dan siswa dengan kemampuan rendah (kelompok bawah). Indikator ini dihitung berdasarkan perbedaan rata-rata skor antara kedua kelompok dan dinyatakan dalam bentuk persentase (DP %). Semakin besar persentase daya pembeda, semakin baik soal tersebut dalam membedakan kemampuan siswa. Berikut adalah analisis beberapa soal berdasarkan daya pembeda mereka:

1. **Soal Nomor 1** memiliki rata-rata skor kelompok atas (Rata2Un) sebesar 3,33 dan kelompok bawah (Rata2As) sebesar 2,00, dengan DP sebesar 33,33% (kategori sedang). Soal ini cukup efektif dalam memilah siswa dengan kemampuan tinggi dan rendah, meskipun daya pembedanya masih bisa ditingkatkan.
2. **Soal Nomor 2** menunjukkan rata-rata skor 4,00 untuk kelompok atas dan 3,33 untuk kelompok bawah, menghasilkan DP sebesar 16,67% (kategori rendah). Soal ini kurang efektif membedakan kemampuan siswa karena selisih rata-rata antar kelompok terlalu kecil, sehingga membutuhkan evaluasi.
3. **Soal Nomor 3** memiliki karakteristik yang sama dengan soal nomor 2, dengan rata-rata skor dan DP yang identik. Daya pembeda rendah menunjukkan soal ini perlu perbaikan agar lebih mampu membedakan kemampuan siswa.
4. **Soal Nomor 4** mencatat rata-rata skor 3,00 untuk kelompok atas dan 2,33 untuk kelompok bawah, dengan DP sebesar 16,67% (kategori rendah). Soal ini juga kurang baik dalam membedakan siswa, mengindikasikan perlunya perbaikan pada soal tersebut.
5. **Soal Nomor 5** memiliki rata-rata skor 4,00 untuk kelompok atas dan 2,33 untuk kelompok bawah, dengan DP sebesar 41,67% (kategori baik). Soal ini merupakan soal terbaik dalam daftar karena daya pembedanya paling tinggi dan efektif dalam membedakan kemampuan siswa.

Daya pembeda sangat penting dalam evaluasi soal, karena membantu menentukan apakah soal tersebut dapat digunakan untuk menilai perbedaan kemampuan siswa secara akurat. Untuk mendapatkan hasil tes yang berkualitas, variasi tingkat kesukaran soal perlu dirancang dengan proporsi

yang tepat, dengan daya pembeda rendah harus diperbaiki atau diganti.

Tabel 3. Daya Pembeda

Dalam penyusunan tes, keberadaan soal-soal dengan tingkat kesukaran yang bervariasi sangat penting untuk menciptakan tes yang efektif. Soal sangat mudah memungkinkan peserta untuk memulai dengan percaya diri, sementara soal mudah memberikan tantangan moderat yang dapat dikerjakan oleh sebagian besar peserta. Soal sedang, di sisi lain, berfungsi untuk menguji kemampuan yang lebih tinggi dan membedakan peserta dengan tingkat pemahaman yang lebih baik. Kombinasi ketiga jenis soal ini memastikan bahwa tes dapat mengevaluasi kemampuan peserta secara menyeluruh, mulai dari pemahaman dasar hingga kemampuan analitis yang lebih mendalam.

Jumlah Subyek= 10		KLP Atas/bawah (n)=3		Butir Soal=5			Un: Unggul As: Assor SB: Simpang Buku		
No	No Btr Asli	Rata2Un	Rata2As	Beda	SB Un	SB As	SB Gab	t	DP (%)
1	1	3,33	2,00	1,33	1,15	0,00	0,67	2,00	33,33
2	2	4,00	3,33	0,67	0,00	1,15	0,67	1,00	16,67
3	3	4,00	3,33	0,67	0,00	0,58	0,33	2,00	16,67
4	4	3,00	2,33	0,67	0,00	0,58	0,33	2,00	16,67
5	5	4,00	2,33	1,67	0,00	0,58	0,33	5,00	41,67

Dengan demikian, pembagian tingkat kesukaran ini tidak hanya membantu mengevaluasi butir soal, tetapi juga memberikan wawasan penting untuk menyempurnakan penyusunan tes agar hasil yang diperoleh dapat mencerminkan kemampuan peserta dengan lebih akurat.

## KESIMPULAN

Penelitian ini mengungkapkan bahwa kualitas butir soal evaluasi materi bilangan cacah kelas 4 SD dapat diukur melalui analisis kuantitatif menggunakan software Anates. Instrumen soal dianalisis berdasarkan aspek validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Hasil analisis menunjukkan bahwa sebagian besar butir soal memenuhi kriteria validitas dan reliabilitas yang baik, dengan koefisien reliabilitas mendekati kategori tinggi ( $\geq 0,7$ ). Tingkat kesukaran soal bervariasi, mulai dari sangat mudah hingga sedang, dengan mayoritas soal tergolong kategori mudah. Namun, beberapa soal memiliki daya pembeda yang rendah, menunjukkan perlunya revisi agar lebih efektif dalam membedakan kemampuan siswa dengan tingkat yang berbeda.

Dengan demikian, penelitian ini memberikan gambaran bahwa soal evaluasi yang telah dianalisis dapat digunakan untuk mengukur pemahaman siswa dengan baik, namun memerlukan beberapa

perbaikan pada butir soal tertentu. Penelitian ini juga diharapkan menjadi acuan bagi penyusunan instrumen evaluasi yang lebih berkualitas pada jenjang pendidikan dasar.

#### **DAFTAR PUSTAKA**

- Akhmadi, M. N. (2021). Analisis butir soal evaluasi tema 1 kelas 4 sdn plumbungan menggunakan program anates. *Ed-Humanistics: Jurnal Ilmu Pendidikan*, 799-806.
- Aqila, D. S. (2022). Analisis Kecocokan Soal Matematika Menggunakan Model Rasch untuk Menilai Tingkat Kesukaran Soal Kelas 5 SD. *Indonesia journal of elementati education*, 1.
- Arif, M. (2015). Penerapan Aplikasi Anates Bentuk Soal Pilihan Ganda. *Eductic. cientific Journal of Informatics Education*, 1-9.
- Fajriah, N. U. (2020). Analisis kemampuan representasi matematis siswa pada materi statistika. *Journal of Educational Review and Research*, Vol. 3 No. 1. Retrieved from STKIP Singkawang.
- Fietri, W. A. (2021). Penelitian Bidang IPA dan Pendidikan IPA Analisis Butir Soal Biologi Kelas XI Madrasah Aliyah Sakinah Kerinci Menggunakan Program Komputer Anates 4.0 For Windows. *Natural Science: , Fietri, W. A. (2021). Natural Science: Jurnal Penelitian Bidang IPA dan Pendidikan IPA Analisis Buti 7(1), 2477-6181.*
- Fiska, J. M. (2021). Analisis Butir Soal Ulangan Harian Ipa Menggunakan Software Anates Pada Pendekatan Teori Tes Klasik. *Natural Science Education Research*,.
- Ghozali, I. (2012). *Aplikasi Analisis Multivariate Dengan Program IBM SPSS 26* . Badan Penerbit Universitas Diponegoro, 1-15.
- mustaqim, M. &. (2024). Analisis butir soal pas matematika peminatan: daya pembeda, tingkat kesukaran, dan kualitas pengecoh. *Al-'Adad: Jurnal Tadris matematik*, 2(1), 2.
- Prastika, Y. D. (2021). Pengaruh Validitas, Reliabilitas dan Tingkat Kesukaran Terhadap Kualitas Butir Soal Ekonomi Menggunakan Software Anates di SMKN 3 Bangkalan. *Prastika, Y. D. (2021). Pengaruh Validitas, Reliabilitas dan Tingkat Kesukaran TSTKIP PGRI Bangkalan*, 1-11.
- Putri, D. &. (2022). Kemampuan pemecahan masalah matematis siswa dalam menyelesaikan soal PISA pada konten bilangan. *Jurnal Penelitian Pembelajaran Matematika*, Vol. 15 No. 1. .
- Suparman, A. ( (2020)). Analisis Distribusi Tingkat Kesulitan Soal dalam Ujian Evaluasi Belajar. *Jurnal Pendidikan*, 15(2), 45-58