

Is wahyuni Listyowati<sup>1</sup>

## ANALISIS BUTIR SOAL MULTIPLE CHOICE MATA PELAJARAN BAHASA INDONESIA SISWA KELAS 4 DI SDN BARENG 5 JOMBANG

### Abstrak

Penelitian ini bertujuan untuk mengembangkan alat ukur evaluasi berupa soal pilihan ganda yang dapat digunakan untuk mengukur pemahaman konsep pembelajaran Bahasa Indonesia pada siswa SDN Bareng 5 Jombang. Pengembangan alat ukur ini melibatkan analisis validitas, reliabilitas, indeks kesulitan, dan daya pembeda soal. Hasil analisis menunjukkan bahwa seluruh soal yang diujikan dinyatakan valid dan reliabel, sehingga dapat digunakan sebagai alat evaluasi yang berkualitas. Uji indeks kesulitan menunjukkan bahwa 10% soal berada pada kategori mudah, sementara 90% soal berada pada kategori sedang, yang menunjukkan tingkat kesulitan yang sesuai dengan kemampuan siswa. Selain itu, hasil uji daya pembeda soal menunjukkan bahwa 67% soal berada pada kategori "baik sekali", dan 33% soal berada pada kategori "baik", yang menunjukkan bahwa soal tersebut efektif dalam membedakan kemampuan siswa. Berdasarkan temuan ini, alat ukur evaluasi yang dikembangkan terbukti layak digunakan untuk mengukur pemahaman konsep Bahasa Indonesia pada siswa SDN Bareng 5 Jombang. Implikasi Penelitian adalah bagi Pengembangan soal pilihan ganda yang valid dan reliabel ini dapat digunakan sebagai model untuk membuat soal evaluasi lainnya dalam pembelajaran Bahasa Indonesia. Soal yang baik, dengan tingkat kesulitan yang tepat dan daya pembeda yang tinggi, dapat meningkatkan kualitas evaluasi dan memberikan gambaran yang akurat mengenai pemahaman siswa terhadap materi yang diajarkan. Secara keseluruhan, hasil penelitian ini menegaskan pentingnya penggunaan soal yang valid, reliabel, dan efektif dalam mengevaluasi pemahaman siswa, yang dapat memberikan kontribusi terhadap peningkatan kualitas pendidikan dan pembelajaran di sekolah.

**Kata kunci:** Analisis Butir Soal, Multiple Choice, Bahasa Indonesia

### Abstract

This study aims to develop an evaluation measuring instrument in the form of multiple-choice questions that can be used to measure the understanding of Indonesian language learning concepts in students of SDN Bareng 5 Jombang. The development of this measuring instrument includes analysis of validity, reliability, difficulty index, and question discrimination power. The results of the analysis show that all questions tested are declared valid and reliable, so they can be used as a quality evaluation tool. The difficulty index test shows that 10% of the questions are in the easy category, while 90% of the questions are in the moderate category, which indicates a level of difficulty that is in accordance with students' abilities. In addition, the results of the question discrimination test show that 67% of the questions are in the "very good" category, and 33% of the questions are in the "good" category, which indicates that the questions are effective in differentiating students' abilities. Based on these findings, the evaluation measuring instrument developed is proven to be feasible to be used to measure the understanding of Indonesian language concepts in students of SDN Bareng 5 Jombang. The implication of the research is that the development of valid and reliable multiple-choice questions can be used as a model for creating other evaluation questions in Indonesian language learning. Good questions, with the right level of difficulty and high discriminating power, can

<sup>1</sup> Sekolah Pascasarjana Universitas PGRI Adi Buana Surabaya  
email: iswahyuni.listyowati70@gmail.com

improve the quality of evaluation and provide an accurate picture of students' understanding of the material being taught. Overall, the results of this study emphasize the importance of using valid, reliable, and effective questions in spreading students' understanding, which can contribute to improving the quality of education and learning in schools.

**Keywords:** Item Analysis, Multiple-Choice, Indonesian

## PENDAHULUAN

Pendidikan merupakan faktor penting dalam membentuk kualitas sumber daya manusia yang berkompeten, salah satunya melalui evaluasi pembelajaran yang dilakukan oleh pendidik. Penilaian terhadap hasil belajar siswa menjadi salah satu metode untuk mengukur sejauh mana kompetensi yang telah dikuasai oleh siswa. Salah satu bentuk evaluasi yang umum digunakan dalam sistem pendidikan adalah ujian dengan soal-soal yang berupa pilihan ganda (multiple choice). Soal pilihan ganda dipilih karena dianggap mampu mengukur pemahaman siswa secara efisien, memungkinkan penilaian atas berbagai materi dengan waktu yang terbatas (Saely & Shaleh, 2023; Tanjung et al., 2024; Unknown, 2024).

Namun, kualitas soal yang digunakan dalam evaluasi sangat berpengaruh terhadap hasil yang diperoleh siswa. Soal yang baik haruslah mencakup indikator kompetensi yang relevan dan mampu mengukur keterampilan serta pengetahuan yang telah dipelajari oleh siswa. Penulis menyadari bahwa pentingnya analisis terhadap butir soal agar dapat diketahui kualitas dan efektivitas soal tersebut dalam mengukur kemampuan siswa. Hal ini berfungsi untuk meningkatkan kualitas pembelajaran dan evaluasi yang dilakukan (Lestari et al., 2023; Zalfa Zuhri et al., 2024).

Fenomena ini menunjukkan bahwa tidak semua guru terlatih atau memiliki pengalaman yang cukup dalam melakukan analisis soal. Beberapa guru mungkin kurang memiliki pemahaman atau keterampilan dalam menilai kualitas soal, sehingga soal yang diberikan kepada siswa bisa jadi kurang efektif atau tidak sesuai dengan tujuan pembelajaran yang diinginkan. Hal ini dapat berdampak pada kualitas evaluasi dan akurasi penilaian terhadap kemampuan siswa. Jika guru tidak memiliki kompetensi dalam menganalisis butir soal, maka soal-soal yang digunakan dalam evaluasi mungkin tidak dapat menggambarkan kemampuan siswa dengan akurat. Misalnya, soal bisa terlalu mudah atau terlalu sulit, atau soal tersebut mungkin tidak efektif dalam membedakan siswa yang benar-benar menguasai materi dengan yang tidak. Akibatnya, hasil ujian bisa menjadi tidak representatif dan mengurangi efektivitas pembelajaran (Mahda et al., 2023; UMAMI, 2022).

Kompetensi guru melakukan analisis soal di sini merujuk pada keterampilan, pengetahuan, dan kemampuan yang diperlukan oleh guru untuk dapat mengevaluasi kualitas soal ujian dengan baik. Hal ini mencakup pemahaman tentang cara menyusun soal yang baik, menentukan tingkat kesulitan soal, serta kemampuan untuk menilai apakah soal tersebut efektif dalam mengukur kompetensi yang diinginkan. Analisis soal adalah proses penting dalam menilai apakah soal yang diajukan sudah mencerminkan tujuan pembelajaran dan kemampuan yang ingin diukur. Ini juga melibatkan evaluasi terhadap aspek-aspek soal seperti daya pembeda (kemampuan soal untuk membedakan siswa yang memahami materi dengan yang tidak), kejelasan soal, kualitas distraktor (jawaban yang salah), dan relevansi soal dengan kurikulum (Angelina et al., 2022; Masulili et al., 2022).

Penelitian ini dilakukan di SDN Bareng 5 Jombang pada mata pelajaran Bahasa Indonesia untuk siswa kelas 4. Sebagai subjek yang memiliki perkembangan kognitif dan bahasa yang sedang berkembang, siswa diharapkan mampu memahami berbagai konsep dasar dalam Bahasa Indonesia, seperti pemahaman bacaan, tata bahasa, dan kosa kata. Dengan demikian, analisis terhadap butir soal pilihan ganda yang digunakan dalam ujian Bahasa Indonesia menjadi sangat penting untuk menilai apakah soal-soal tersebut sesuai dengan tujuan pembelajaran dan mampu menggambarkan capaian kompetensi siswa dengan akurat.

Melalui penelitian ini, diharapkan dapat diperoleh gambaran tentang kualitas soal pilihan ganda pada ujian Bahasa Indonesia di SDN Bareng 5 Jombang, serta mengetahui apakah soal-soal tersebut telah sesuai dengan standar pendidikan yang diharapkan dan efektif dalam mengukur kemampuan siswa. Dengan demikian, hasil analisis ini dapat memberikan rekomendasi bagi pengembangan soal-soal evaluasi di masa mendatang dan perbaikan dalam proses pembelajaran Bahasa Indonesia.

## METODE

### 1. Jenis Penelitian

Penelitian ini menggunakan pendekatan kuantitatif deskriptif. Penelitian bertujuan untuk menganalisis kualitas butir soal multiple choice pada mata pelajaran Bahasa Indonesia yang diberikan kepada siswa kelas 4 di SDN Bareng 5 Jombang. Fokus utama adalah untuk mengevaluasi efektivitas soal dalam mengukur kompetensi siswa dan menilai berbagai aspek teknis soal seperti tingkat kesulitan, daya pembeda, reliabilitas, dan kualitas distraktor.

### 2. Pendekatan Penelitian

Pendekatan yang digunakan dalam penelitian ini adalah analisis butir soal. Soal yang dianalisis adalah soal ujian pilihan ganda yang digunakan dalam evaluasi mata pelajaran Bahasa Indonesia untuk siswa kelas 4. Analisis dilakukan secara kuantitatif dengan menggunakan berbagai metode statistik untuk mengukur kualitas soal.

### 3. Populasi dan Sampel

Populasi: Semua soal ujian pilihan ganda dalam mata pelajaran Bahasa Indonesia yang diberikan kepada siswa kelas 4 SDN Bareng 5 Jombang pada tahun ajaran yang sedang berlangsung. Sampel: Sampel penelitian ini terdiri dari sejumlah soal pilihan ganda yang dipilih secara acak dari ujian Bahasa Indonesia yang dilaksanakan selama periode penelitian. Sampel soal yang dipilih akan mencakup soal dari berbagai materi yang diajarkan, dengan tujuan agar hasil analisis dapat mewakili kualitas soal secara keseluruhan.

### 4. Instrumen Penelitian

Instrumen utama yang digunakan dalam penelitian ini adalah butir soal pilihan ganda dari ujian Bahasa Indonesia yang diberikan kepada siswa kelas 4 SDN Bareng 5 Jombang. Instrumen tambahan untuk menganalisis kualitas soal meliputi: (a) Rumus untuk menghitung tingkat kesulitan soal (misalnya dengan menghitung persentase siswa yang menjawab benar), (b) Rumus untuk menghitung daya pembeda soal (untuk mengetahui apakah soal dapat membedakan siswa yang mampu menguasai materi dengan yang tidak), (c) Koefisien reliabilitas soal untuk mengukur konsistensi hasil ujian, dan (d) Analisis distraktor untuk mengevaluasi kualitas pilihan jawaban yang salah dalam soal.

### 5. Prosedur Penelitian

Tahap 1: Pengumpulan Data, Data dikumpulkan dengan mengumpulkan semua soal pilihan ganda yang digunakan pada ujian Bahasa Indonesia untuk siswa kelas 4 di SDN Bareng 5 Jombang. Soal yang terkumpul kemudian diseleksi untuk memastikan bahwa soal tersebut sesuai dengan kriteria penelitian.

Tahap 2: Pengolahan dan Analisis Data Setelah data soal terkumpul, setiap soal akan dianalisis menggunakan teknik statistik sebagai berikut: (a) uji validitas soal, (b) Reliabilitas Soal: Menggunakan koefisien reliabilitas untuk mengukur konsistensi soal dalam mengukur kemampuan siswa pada ujian yang berbeda, (c) Tingkat Kesulitan Soal: Dihitung dengan membagi jumlah siswa yang menjawab benar dengan total jumlah siswa yang mengikuti ujian, kemudian dikategorikan menjadi soal yang mudah, sedang, atau sulit, (d) Daya Pembeda Soal: Menghitung seberapa baik soal dapat membedakan antara siswa yang memiliki pemahaman materi yang baik dengan siswa yang kurang memahami materi.

Tahap 3: Penyusunan Laporan Setelah analisis data dilakukan, hasil penelitian akan disusun dalam bentuk laporan yang mencakup kesimpulan dan rekomendasi terkait kualitas soal dan kemungkinan perbaikan di masa depan.

## HASIL DAN PEMBAHASAN

### Hasil Penelitian

Uji validitas dilakukan dengan dua cara yaitu (1) membandingkan nilai pearson correlation atau nilai rhitung dengan nilai rtabel, nilai rhitung harus lebih besar dari nilai rtabel ( $rhitung > rtabel$ ). (2) Mengkombinasikan nilai pearson correlation dengan signifikansi yang diperoleh, nilai yang diperoleh harus positif dan didukung nilai signifikansi  $<0.05$ . Hasil pengujian dengan menggunakan pearson correlation dapat dilihat pada tabel 1.

Tabel 1. Hasil uji validitas

No Soal	Hasil R Hitung	Nilai Signifikan	Validitas
---------	----------------	------------------	-----------

item01	.618**	0.000	valid
item02	.973**	0.000	valid
item03	.739**	0.000	valid
item04	.973**	0.000	valid
item05	.973**	0.000	valid
item06	.590**	0.001	valid
item07	.557**	0.001	valid
item08	.428*	0.018	valid
item09	.973**	0.000	valid
item10	.739**	0.000	valid
item11	.913**	0.000	valid
item12	.902**	0.000	valid
item13	.498**	0.005	valid
item14	.767**	0.000	valid
item15	.907**	0.000	valid
item16	.739**	0.000	valid
item17	.538**	0.002	valid
item18	.973**	0.000	valid
item19	.767**	0.000	valid
item20	.634**	0.000	valid
item21	.973**	0.000	valid
item22	.938**	0.000	valid
item23	.849**	0.000	valid
item24	.767**	0.000	valid
item25	.849**	0.000	valid
item26	.739**	0.000	valid
item27	.538**	0.002	valid
item28	.444*	0.014	valid
item29	.767**	0.000	valid
item30	.634**	0.000	valid

Pada penelitian ini berdasarkan data diatas dapat dinyatakan bahwa uji validitas soal dikategorikan valid dengan nilai tinggi sebesar 973 dan nilai signifikan paling tinggi adalah 0.000 sehingga instrumen tersebut layak digunakan sebagai alat ukur dalam mengevaluasi pembelajaran Bahasa Indonesia di SDN Bareng 5 Jombang.

Uji reliabilitas berdasarkan Chronbach alpha membandingkan antara nilai hasil uji dengan besaran 0.06. Apabila nilai hasil uji lebih besar dari 0.06 maka butir soal dinyatakan reliabel. Data ditampilkan yang dinyatakan dalam tabel 2.

Tabel 2. Uji Reliabilitas

**Reliability Statistics**

Cronbach's Alpha	N of Items
.975	30

Parameter kelayakan suatu instrumen tidak hanya ditentukan oleh validitas dan reliabilitas namun juga harus dapat menggambarkan perbedaan pengetahuan dan kemampuan peserta didik oleh karena itu dibutuhkan instrumen indeks kesukaran dan daya pembeda soal untuk dapat membedakan tingkat kemampuan masing-masing peserta didik (Ismiyati et al., 2023; Rahayu et al., 2023; Sari, 2022).

Indeks kesukaran butir soal dilakukan untuk mengidentifikasi tingkat kesukaran masing-masing butir soal dengan membandingkan hasil uji dan indeks kesukaran butir soal pada tabel 3. Hasil uji indeks kesukaran menunjukkan butir soal berada pada di kategori mudah 10%, dan butir soal berada pada di kategori sedang 90%. Hasil lebih rinci dipaparkan pada tabel 3.

Tabel 3. Indeks kesukaran butir soal

No Soal	Nilai R Hitung	Kriteria Pengambilan Keputusan	Kriteria
1	0.7000		Sedang
2	0.4333		Sedang
3	0.7000		Sedang
4	0.4333		Sedang
5	0.4333		Sedang
6	0.6000		Sedang
7	0.8000		mudah
8	0.8667		mudah
9	0.4333		Sedang
10	0.7000		Sedang
11	0.4000		Sedang
12	0.5000		Sedang
13	0.7000	Indeks tingkat kesukaran 0.00-0,15 =sangat sukar 0,16-0,30 = sukar 0,31-0,70 = sedang 0,71-0,85 = mudah 0,86-1,00= sangat mudah	Sedang
14	0.6333		Sedang
15	0.4667		Sedang
16	0.7000		Sedang
17	0.6333		Sedang
18	0.4333		Sedang
19	0.6333		Sedang
20	0.6333		Sedang
21	0.4333		Sedang
22	0.4667		Sedang
23	0.4333		Sedang
24	0.6333		Sedang
25	0.4333		Sedang
26	0.7000		Sedang
27	0.6333		Sedang
28	0.7667		mudah
29	0.6333		Sedang
30	0.6333		Sedang

Uji daya pembeda soal berfungsi untuk menggambarkan perbedaan kemampuan yang dimiliki oleh masing-masing responden dan membandingkannya dengan kategori daya pembeda soal. Hasil uji menunjukkan butir soal berada pada level baik sekali dengan presentasi 67%, dan soal berada di kategori baik dengan presentasi 33%. Hasil lebih rinci dipaparkan pada tabel 4.

Table 4. Hasil uji daya pembeda soal

No Soal	Hasil R Hitung	Kriteria interpretasi	Validitas
1	0.591		baik
2	0.970		baik sekali
3	0.719		baik sekali
4	0.970		baik sekali
5	0.970		baik sekali
6	0.559		baik
7	0.530		baik
8	0.401		baik
9	0.970	0.70-1.00 = baik sekali 0.40-0.69 = baik 0.20-0.39=cukup 0.00-0.19=jelek	baik sekali
10	0.719		baik sekali
11	0.905		baik sekali
12	0.892		baik sekali
13	0.465		baik
14	0.748		baik sekali
15	0.899		baik sekali

16	0.719	baik sekali
17	0.506	baik
18	0.970	baik sekali
19	0.748	baik sekali
20	0.606	baik
21	0.970	baik sekali
22	0.932	baik sekali
23	0.835	baik sekali
24	0.748	baik sekali
25	0.835	baik sekali
26	0.719	baik sekali
27	0.506	baik
28	0.412	baik
29	0.748	baik sekali
30	0.606	baik

Pada penelitian ini diperoleh hasil uji daya pembeda soal menyatakan bahwa seluruh soal yang diujikan dalam kategori baik sekali dan baik. Sehingga dapat dinyatakan bahwa keseluruhan butir soal memiliki kualitas yang baik sehingga tidak memerlukan perbaikan dan bisa digunakan sebagai alat evaluasi dalam pembelajaran.

### Pembahasan

Berdasarkan data yang diperoleh, uji validitas instrumen evaluasi pembelajaran Bahasa Indonesia di SDN Bareng 5 Jombang menunjukkan nilai validitas tinggi sebesar 0,973 dengan tingkat signifikansi 0,000. Nilai ini mengindikasikan bahwa instrumen tersebut sangat layak digunakan untuk mengevaluasi pembelajaran Bahasa Indonesia di sekolah tersebut. Validitas instrumen adalah ukuran yang menunjukkan sejauh mana alat ukur dapat mengukur apa yang seharusnya diukur. Nilai validitas yang tinggi menunjukkan bahwa instrumen tersebut efektif dalam mengukur kompetensi yang ditargetkan. Beberapa penelitian sebelumnya mendukung pentingnya validitas instrumen dalam evaluasi pembelajaran.

Penelitian ini didukung oleh (Ernawati, 2023; Marambaawang et al., 2023; Yudhanto, 2022) menyatakan Penelitian mereka menunjukkan bahwa instrumen evaluasi pembelajaran yang memiliki validitas tinggi dapat secara efektif mengukur pemahaman siswa terhadap materi pelajaran. Sejalan dengan penelitian Safitri dkk. (2023) menyatakan Studi ini menekankan bahwa instrumen evaluasi dengan tingkat validitas yang tinggi berkontribusi signifikan terhadap peningkatan hasil belajar siswa. Chotimah (2024) Penelitian ini menyoroti bahwa soal yang valid dan reliabel dapat digunakan sebagai alat evaluasi yang berkualitas, sehingga meningkatkan efektivitas pembelajaran.

Nilai signifikansi 0,000 yang diperoleh dalam penelitian ini menunjukkan bahwa hubungan antara instrumen evaluasi dan hasil belajar siswa sangat kuat dan tidak terjadi secara kebetulan. Hal ini menegaskan bahwa instrumen yang digunakan memiliki kontribusi signifikan terhadap peningkatan kualitas evaluasi pembelajaran di SDN Bareng 5 Jombang.

Berdasarkan hasil analisis indeks kesukaran butir soal, diperoleh bahwa 10% soal berada pada kategori mudah dan 90% pada kategori sedang. Analisis tingkat kesukaran ini penting untuk menilai kualitas dan efektivitas evaluasi pembelajaran. Tingkat kesukaran soal yang ideal sebaiknya seimbang, dengan proporsi soal mudah, sedang, dan sulit yang sesuai. Soal yang terlalu mudah tidak dapat membedakan kemampuan siswa secara efektif, sedangkan soal yang terlalu sulit dapat menurunkan motivasi belajar siswa.

Penelitian ini didukung oleh (Jumini et al., 2023; Kurniawan, 2022; Nurul Muchlizani A et al., 2023) Penelitian ini menganalisis tingkat kesukaran, dan daya pembeda pada soal ulangan tengah semester mata pelajaran Bahasa Indonesia kelas IV SD Negeri Waruruma. Hasilnya menunjukkan bahwa mayoritas soal berada pada kategori mudah, dengan 60% soal mudah dan 40% sedang. Penelitian ini menekankan pentingnya keseimbangan tingkat kesukaran soal untuk memastikan evaluasi yang efektif. Sejalan dengan penelitian Fitriyani (2023) Studi ini menganalisis kualitas butir soal pada Ulangan Akhir Semester Bahasa Indonesia di SDN 1 Ciparay. Ditemukan bahwa 32% soal sangat mudah, 32% mudah, dan 34% sedang. Penelitian ini menyoroti bahwa distribusi soal yang seimbang antara mudah dan sedang dapat membantu

dalam mengukur kemampuan siswa secara lebih akurat. Azis (2024) Penelitian ini menganalisis butir soal UAS Gasal 2017/2018 pada mata pelajaran Bahasa Indonesia kelas IV SD. Hasilnya menunjukkan bahwa sebagian besar soal memiliki tingkat kesukaran sedang, dengan rentang -2,51 hingga 1,46. Studi ini menekankan bahwa soal dengan tingkat kesukaran sedang lebih efektif dalam mengevaluasi pemahaman siswa tanpa menyebabkan kebingungan atau kejemuhan.

Berdasarkan hasil analisis daya pembeda soal dalam penelitian ini, ditemukan bahwa 67% butir soal berada pada kategori "baik sekali" dan 33% pada kategori "baik". Daya pembeda yang tinggi menunjukkan bahwa soal-soal tersebut efektif dalam membedakan kemampuan siswa yang memiliki pemahaman baik dari yang kurang memahami materi. Daya pembeda soal adalah indikator penting dalam evaluasi pembelajaran, karena mencerminkan sejauh mana soal dapat membedakan antara siswa yang berprestasi tinggi dan rendah. Soal dengan daya pembeda yang baik dapat mengidentifikasi siswa yang benar-benar menguasai materi, sehingga hasil evaluasi lebih akurat dan dapat dipercaya.

Penelitian ini didukung oleh (Aprilia Ningrum, 2023; Hanan et al., 2023; Nur Cahyo et al., 2022) Dalam penelitian mereka tentang analisis tingkat kesukaran dan daya pembeda pada butir tes soal ujian tengah semester Bahasa Indonesia, ditemukan bahwa mayoritas soal memiliki daya pembeda yang baik, yang menunjukkan bahwa soal tersebut efektif dalam membedakan kemampuan siswa. Sejalan dengan penelitian Hidayanti (2021) Penelitiannya mengenai analisis butir soal Bahasa Indonesia pada evaluasi pembelajaran menunjukkan bahwa soal dengan daya pembeda tinggi dapat meningkatkan kualitas evaluasi dan membantu dalam mengidentifikasi area yang perlu perbaikan dalam pembelajaran. Sari (2023) Studi tentang analisis butir soal pilihan ganda berorientasi Higher Order Thinking Skill (HOTS) pada tema tertentu menunjukkan bahwa soal dengan daya pembeda baik dapat meningkatkan kualitas evaluasi dan membantu dalam mengidentifikasi area yang perlu perbaikan dalam pembelajaran.

Hasil analisis daya pembeda soal dalam penelitian ini menunjukkan bahwa mayoritas soal memiliki daya pembeda yang tinggi, yang berarti soal-soal tersebut efektif dalam membedakan kemampuan siswa. Hal ini sejalan dengan temuan dari penelitian sebelumnya yang menekankan bahwa soal dengan daya pembeda baik dapat meningkatkan kualitas evaluasi dan membantu dalam mengidentifikasi area yang perlu perbaikan dalam pembelajaran. Oleh karena itu, disarankan untuk terus mempertahankan dan meningkatkan kualitas soal dengan memastikan bahwa soal memiliki daya pembeda yang baik, sehingga evaluasi pembelajaran dapat mencerminkan kemampuan siswa secara akurat.

## SIMPULAN

Berdasarkan hasil penelitian ini, dapat disimpulkan bahwa pengembangan alat ukur evaluasi berupa soal pilihan ganda untuk mengukur pemahaman konsep tentang pembelajaran Bahasa Indonesia di SDN Bareng 5 Jombang telah berhasil. Analisis terhadap soal menunjukkan hasil yang sangat positif, dengan seluruh soal yang diuji dinyatakan valid dan reliabel. Ini berarti soal-soal yang digunakan telah memenuhi kriteria untuk digunakan sebagai alat evaluasi yang berkualitas dalam pembelajaran Bahasa Indonesia. Hasil uji indeks kesulitan soal menunjukkan bahwa mayoritas soal berada dalam kategori sedang (90%), yang berarti soal tersebut memiliki tingkat kesulitan yang sesuai dengan kemampuan siswa. Hanya 10% soal yang termasuk dalam kategori mudah. Selain itu, hasil uji daya pembeda menunjukkan bahwa 67% soal berada pada kategori "baik sekali", dan 33% soal berada pada kategori "baik", yang menandakan bahwa soal tersebut efektif dalam membedakan kemampuan siswa yang berprestasi tinggi dan rendah. Implikasi Penelitian adalah bagi Pengembangan soal pilihan ganda yang valid dan reliabel ini dapat digunakan sebagai model untuk membuat soal evaluasi lainnya dalam pembelajaran Bahasa Indonesia. Soal yang baik, dengan tingkat kesulitan yang tepat dan daya pembeda yang tinggi, dapat meningkatkan kualitas evaluasi dan memberikan gambaran yang akurat mengenai pemahaman siswa terhadap materi yang diajarkan. Secara keseluruhan, hasil penelitian ini menegaskan pentingnya penggunaan soal yang valid, reliabel, dan efektif dalam mengevaluasi pemahaman siswa, yang dapat memberikan kontribusi terhadap peningkatan kualitas pendidikan dan pembelajaran di sekolah.

**DAFTAR PUSTAKA**

- Angelina, G. M., Elfitra, L., & Zaitun, Z. (2022). Analisis Butir Soal Menggunakan Teori Pengukuran Klasik Dan Taksonomi Bloom Pada Ujian Akhir Semester Mata Pelajaran Bahasa Indonesia Kelas VII Sekolah .... Student Online Journal (Soj) ....
- Aprilia Ningrum, W. (2023). Analisis Butir Soal Ulangan Harian Pembelajaran Ipa Pada Kelas VIII Di SMPN 21 Semarang. *Educatio*. [Https://Doi.Org/10.29408/Edc.V18i1.18291](https://doi.org/10.29408/edc.v18i1.18291)
- Ernawati, E. (2023). Analisis Soal Ujian Matematika Berkategori Higher Order Thinking Skills (Hots) Siswa Smp Dengan Analisis Rasch Model. *Jurnal Didaktika Pendidikan Dasar*, 7(1). [Https://Doi.Org/10.26811/Didaktika.V7i1.1027](https://doi.org/10.26811/didaktika.v7i1.1027)
- Hanan, M. P., Jannah, R. R., & Alim, J. A. (2023). Analisis Butir Soal Matematika Materi Kpk Dan Fpb Berbasis Cerita Di Sdn 111 Pekanbaru. *Journal Of Educational Learning And Innovation (Elia)*, 3(1). [Https://Doi.Org/10.46229/Elia.V3i1.538](https://doi.org/10.46229/elia.v3i1.538)
- Ismyati, I., Raharjo, T. H., Tusyanah, T., & Sholikah, M. (2023). Pelatihan Analisis Butir Soal Berdasarkan Teori Tes Klasik Berbantuan Iteman Untuk Meningkatkan Kualitas Instrumen Penilaian. *Japi (Jurnal Akses Pengabdian Indonesia)*, 8(2). [Https://Doi.Org/10.33366/Japi.V8i2.5064](https://doi.org/10.33366/japi.v8i2.5064)
- Jumini, S., Madnasri, S., Cahyono, E., & Parmin, P. (2023). Analisis Kualitas Butir Soal Pengukuran Literasi Sains Melalui Teori Tes Klasik Dan Rasch Model. Prosiding Seminar Nasional Pascasarjana Unnes.
- Kurniawan, T. (2022). Analisis Butir Soal Ulangan Akhir Semester Gasal Mata Pelajaran Ips Sekolah Dasar. *Jurnal Basicedu*, 6(4).
- Lestari, A. S., Fitrianna, A. Y., & Zanthy, L. S. (2023). Analisis Butir Soal Tes Materi Sistem Persamaan Linear Dua Variabel Pada Siswa Kelas VIII. *Jurnal Pembelajaran Matematika Inovatif*, 6(1).
- Mahda, A., Arfiyanti, M. P., Novitasari, A., & Romadhoni, R. (2023). Analisis Deskriptif Kualitas Soal Multiple Choice Questions (Mcq) Mini Kuis Tutorial Di Fakultas Kedokteran Universitas Muhammadiyah Semarang. *Jurnal Ilmu Kedokteran Dan Kesehatan*, 10(6). [Https://Doi.Org/10.33024/Jikk.V10i6.9932](https://doi.org/10.33024/jikk.v10i6.9932)
- Marambaawang, D., Bano, V. O., & Hada Enda, R. R. (2023). Analisis Kualitas Butir Soal Penilaian Akhir Semester Gasal Tahun 2021/2022 Menggunakan Iteman Di Smp Negeri 1 Kambera. *Dharmas Education Journal (De\_Journal)*, 4(1). [Https://Doi.Org/10.56667/Dejournal.V4i1.961](https://doi.org/10.56667/dejournal.v4i1.961)
- Masulili, R. R., Dama, L., & Abdul, A. (2022). Analisis Butir Soal Semester Genap Mata Pelajaran Biologi Kelas XI Di SMAN 3 Gorontalo Dan SMAN 5 Gorontalo. *Jambura Edu Biosfer Journal*, 3(2). [Https://Doi.Org/10.34312/Jebj.V3i2.7211](https://doi.org/10.34312/jebj.v3i2.7211)
- Nur Cahyo, A., Luriawati, D., & Wagiran, W. (2022). Analisis Butir Soal Penilaian Keterampilan Kebahasaan Pada Pembelajaran Teks Eksplanasi Kelas XI. *Jbsi: Jurnal Bahasa Dan Sastra Indonesia*, 2(01). [Https://Doi.Org/10.47709/Jbsi.V2i01.1493](https://doi.org/10.47709/jbsi.v2i01.1493)
- Nurul Muchlizani A, Sitti Mania, & Muhammad Nur Akbar Rasyid. (2023). Analisis Kualitas Butir Soal Ujian Akhir Semester Mata Pelajaran Akidah Akhlak Kelas V MI Radhiyatul Adawiyah Makassar. *Inspiratif Pendidikan*, 12(1). [Https://Doi.Org/10.24252/Ip.V12i1.32040](https://doi.org/10.24252/Ip.V12i1.32040)
- Rahayu, W., Sulaeman, E., Arnawisuda Ningsi, B., Arofah, I., & Akbari, W. (2023). Analisis Kemampuan Pemahaman Guru Sekolah Dasar Dalam Mengembangkan Butir Soal Berbasis Literasi Numerik. *Perduli: Jurnal Pengabdian Kepada Masyarakat*, 4(01). [Https://Doi.Org/10.21009/Perduli.V4i01.30490](https://doi.org/10.21009/Perduli.V4i01.30490)
- Saely, E., & Shaleh, S. (2023). Analisis Butir Soal Penilaian Akhir Semester (Pas) Pada Mata Pelajaran Ppkn Menggunakan Aplikasi Winstep Dan Item And Tes Analysis (Iteman) 4.3. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, 8(3).
- Sari, R. P. (2022). Analisis Butir Soal Ujian Akhir Semester Genap Mata Pelajaran Matematika Kelas XI SMK Negeri 1 Percut Sei Tuan T.A. 2019/2020. *Jurnal Penelitian Pendidikan Mipa*, 7(1). [Https://Doi.Org/10.32696/Jp2mipa.V7i1.1360](https://doi.org/10.32696/jp2mipa.v7i1.1360)
- Tanjung, M. A. H. R., Fahmi, A. A., Rahmanita, F., Habibah, I. F., & Qomari, N. (2024). Analisis Butir Soal Penilaian Akhir Tahun Pelajaran Bahasa Arab Kelas VII Mts Al-Ma'arif Rakit Banjarnegara Jawa Tengah. *Mantiqu Tayr: Journal Of Arabic Language*, 4(1). [Https://Doi.Org/10.25217/Mantiquayr.V4i1.4038](https://doi.org/10.25217/mantiquayr.v4i1.4038)
- Umami, F. (2022). Analisis Butir Soal Tes Pilihan Ganda Ujian Nasional Tahun Pelajaran 2018-

- 2019 Mata Pelajaran Bahasa Indonesia Jenjang Sltp (Sekolah Lanjut Tingkat Pertama). Repository.Uinjkt.Ac.Id.
- Unknown. (2024). Analisis Butir Soal Dengan Menggunakan Spss. Masih Belajar.
- Yudhanto, Y. (2022). Rancang Bangun Aplikasi Exam Dengan Metode Analisis Butir Soal. Indonesian Journal Of Applied Informatics, 5(2). <Https://Doi.Org/10.20961/Ijai.V5i2.43505>
- Zalfa Zuhri, N., Sopian, A., Sofyan Sauri, & Nurbayan, Y. (2024). Analisis Validitas Dan Reliabilitas Soal Bahasa Arab Melalui Website Opexams Pembuat Soal Berbasis Ai. Jurnal Pendidikan Modern, 9(2). <Https://Doi.Org/10.37471/Jpm.V9i2.863>