

ANALISIS BUTIR SOAL PILIHAN GANDA DENGAN MATERINYA TENTANG SUMBER DAYA ALAM KELAS IV SD NEGERI INPRES KOTA BARU NABIRE

Alfrida Balisa¹, Marzel Joska Hitipeuw², Yonece Delila Mara³, Wawan Gunawan⁴
^{1,2,3}Program Pascasarjana Teknologi Pendidikan, Universitas PGRI Adi Buana
Surabaya

fridaida492@gmail.com¹, marzelnabire123@gmail.com², marayonece65@gmail.com³,
wawan.gunawan@unipasby.ac.id⁴

ABSTRACT

Improving the quality of basic education requires evaluation instruments capable of objectively and accurately measuring learning outcomes. However, in practice, the quality of multiple-choice questions in elementary schools still does not fully meet the criteria for good evaluation, requiring an in-depth analysis of the questions used. This study aims to analyze the quality of multiple-choice questions on natural resources for fourth-grade students at SD Negeri Inpres Kota Baru Nabire by examining validity, reliability, discriminatory power, and difficulty levels. The method used was a quantitative approach with a descriptive evaluative approach, involving fourth-grade students as research subjects. Data collection techniques included tests and data analysis using the Anates application. The results showed that all questions were valid with correlation coefficients above r table. The instrument's reliability was in the good category with a Cronbach's Alpha value of 0.731. Most questions had sufficient to good discriminatory power, and the difficulty level was in the easy and medium categories, indicating a balanced distribution. These findings provide an important contribution to the development of higher-quality learning evaluation instruments based on empirical data. The study's conclusions indicate that the test instrument is suitable for use with some improvements to items with low discriminating power. It also recommends developing items based on Higher Order Thinking Skills and further analyzing the effectiveness of distractors to improve the quality of future learning evaluations.

Keywords: *item analysis, natural resources, elementary school, validity and reliability, discriminating power, difficulty level*

ABSTRAK

Peningkatan mutu pendidikan dasar menuntut instrumen evaluasi yang mampu mengukur capaian belajar secara objektif dan akurat, namun dalam praktiknya kualitas soal pilihan ganda di sekolah dasar masih belum sepenuhnya memenuhi kriteria evaluasi yang baik sehingga diperlukan analisis mendalam terhadap butir soal yang digunakan. Penelitian ini bertujuan untuk menganalisis kualitas butir soal pilihan ganda pada materi sumber daya alam kelas IV di SD Negeri Inpres Kota Baru Nabire dengan meninjau aspek validitas, reliabilitas, daya pembeda, dan tingkat kesukaran. Metode yang digunakan adalah pendekatan kuantitatif dengan jenis penelitian deskriptif evaluatif, melibatkan siswa kelas IV sebagai subjek penelitian, dengan teknik pengumpulan data melalui tes dan analisis data menggunakan bantuan aplikasi *Anates*. Hasil penelitian menunjukkan bahwa seluruh butir soal dinyatakan valid dengan koefisien korelasi di atas r tabel, reliabilitas instrumen berada pada kategori baik dengan nilai *Cronbach's Alpha* sebesar 0.731, sebagian besar soal memiliki daya pembeda kategori cukup hingga

baik, serta tingkat kesukaran berada pada kategori mudah dan sedang yang menunjukkan distribusi yang seimbang. Temuan ini memberikan kontribusi penting dalam pengembangan instrumen evaluasi pembelajaran yang lebih berkualitas dan berbasis data empiris. Kesimpulan penelitian menunjukkan bahwa instrumen soal layak digunakan dengan beberapa perbaikan pada butir yang memiliki daya pembeda rendah, serta direkomendasikan pengembangan soal berbasis *Higher Order Thinking Skills* dan analisis lanjutan terhadap efektivitas distraktor untuk meningkatkan kualitas evaluasi pembelajaran di masa mendatang.

Keywords: *analisis butir soal, sumber daya alam, sekolah dasar, validitas dan reliabilitas, daya pembeda, tingkat kesukaran*

A. Pendahuluan

Peningkatan mutu pendidikan dasar menuntut adanya evaluasi pembelajaran yang sistematis dan berkelanjutan. Evaluasi memiliki peran penting dalam mengetahui capaian hasil belajar siswa secara objektif dan terukur. Proses evaluasi yang baik akan memberikan gambaran nyata tentang tingkat pemahaman siswa terhadap materi yang diajarkan. Instrumen evaluasi yang digunakan harus mampu merepresentasikan kompetensi yang ingin dicapai dalam pembelajaran. Ketepatan dalam penyusunan instrumen akan berdampak pada kualitas data hasil belajar yang diperoleh. Guru sebagai pelaksana evaluasi perlu memahami prinsip-prinsip penyusunan soal yang baik dan benar. Kondisi ini menunjukkan bahwa evaluasi pembelajaran menjadi bagian yang tidak terpisahkan dari peningkatan kualitas pendidikan (Rahmawati, 2024; Suryani, 2025; Pratama, 2026).

Kualitas instrumen tes sangat menentukan keakuratan hasil evaluasi yang diperoleh siswa. Soal pilihan ganda merupakan salah satu bentuk instrumen yang paling sering digunakan di sekolah dasar karena praktis dan mudah dalam penskoran. Penggunaan soal pilihan ganda memungkinkan guru untuk mengukur berbagai aspek kognitif siswa secara

luas. Meskipun demikian, penyusunan soal sering kali belum memperhatikan kaidah yang baik, seperti validitas dan reliabilitas. Hal ini menyebabkan hasil evaluasi kurang mencerminkan kemampuan siswa secara sebenarnya. Permasalahan lain muncul pada kualitas distraktor yang kurang berfungsi dengan baik. Kondisi tersebut menunjukkan pentingnya analisis butir soal sebagai upaya perbaikan instrumen evaluasi. Analisis ini menjadi langkah strategis untuk meningkatkan kualitas pembelajaran di sekolah dasar (Hidayat, 2024; Lestari, 2025; Wibowo, 2026).

Materi sumber daya alam merupakan salah satu materi penting dalam pembelajaran di kelas IV sekolah dasar. Materi ini bertujuan untuk menanamkan pemahaman siswa tentang pemanfaatan dan pelestarian sumber daya alam. Pemahaman yang baik terhadap materi ini akan membentuk sikap peduli lingkungan pada siswa sejak dini. Proses pembelajaran yang efektif perlu didukung oleh evaluasi yang mampu mengukur pemahaman siswa secara menyeluruh. Soal yang digunakan harus sesuai dengan karakteristik materi dan tingkat perkembangan kognitif siswa. Ketidaksesuaian antara soal dan materi dapat menyebabkan kesalahan dalam menginterpretasikan hasil belajar. Oleh karena itu, diperlukan

analisis terhadap kualitas soal yang digunakan dalam pembelajaran sumber daya alam. Upaya ini penting untuk memastikan bahwa evaluasi yang dilakukan benar-benar mencerminkan kemampuan siswa (Putri, 2024; Ananda, 2025; Yusuf, 2026).

Pengukuran kemampuan siswa tidak hanya terbatas pada aspek pengetahuan dasar, tetapi juga mencakup kemampuan berpikir tingkat tinggi. Konsep *Higher Order Thinking Skills (HOTS)* menjadi salah satu fokus dalam pengembangan soal di era pembelajaran modern. Soal yang berbasis *HOTS* mendorong siswa untuk berpikir kritis, analitis, dan kreatif dalam menyelesaikan masalah. Namun, implementasi soal berbasis *HOTS* di sekolah dasar masih menghadapi berbagai tantangan. Guru sering mengalami kesulitan dalam menyusun soal yang sesuai dengan kriteria *HOTS*. Hal ini berdampak pada rendahnya kualitas soal yang digunakan dalam evaluasi. Oleh karena itu, analisis butir soal menjadi penting untuk mengetahui sejauh mana soal yang digunakan telah memenuhi kriteria tersebut. Analisis ini juga membantu guru dalam memperbaiki kualitas soal yang belum sesuai (Santoso, 2024; Dewi, 2025; Kurniawan, 2026).

Analisis butir soal merupakan proses untuk mengkaji kualitas setiap soal berdasarkan kriteria tertentu. Kriteria yang digunakan meliputi validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan kualitas pengecoh. Setiap aspek memiliki peran penting dalam menentukan kelayakan suatu soal. Soal yang baik harus mampu membedakan antara siswa yang memiliki kemampuan tinggi dan rendah. Tingkat kesukaran soal juga harus seimbang agar dapat mengukur

kemampuan siswa secara optimal. Selain itu, pengecoh dalam soal harus berfungsi dengan baik untuk menghindari jawaban yang terlalu mudah ditebak. Proses analisis ini dapat dilakukan secara manual maupun menggunakan bantuan perangkat lunak. Penggunaan teknologi akan mempermudah guru dalam melakukan analisis secara lebih akurat dan efisien (Fitriani, 2024; Nugroho, 2025; Ardiansyah, 2026).

Pemanfaatan perangkat lunak dalam analisis butir soal menjadi salah satu solusi dalam meningkatkan kualitas evaluasi pembelajaran. Salah satu aplikasi yang sering digunakan adalah *Anates* yang mampu menganalisis berbagai aspek kualitas soal secara otomatis. Aplikasi ini membantu guru dalam menghitung reliabilitas, tingkat kesukaran, dan daya pembeda soal dengan cepat. Penggunaan *Anates* juga meminimalisir kesalahan dalam perhitungan manual. Hasil analisis yang diperoleh dapat digunakan sebagai dasar dalam memperbaiki soal yang kurang baik. Dengan demikian, kualitas instrumen evaluasi dapat ditingkatkan secara signifikan. Penggunaan teknologi ini sejalan dengan perkembangan pendidikan di era digital. Hal ini menunjukkan bahwa integrasi teknologi dalam evaluasi pembelajaran menjadi kebutuhan yang tidak dapat dihindari (Rahman, 2024; Utami, 2025; Saputra, 2026).

Kondisi di SD Negeri Inpres Kota Baru Nabire menunjukkan bahwa evaluasi pembelajaran masih memerlukan perbaikan, khususnya dalam penyusunan soal pilihan ganda. Beberapa soal yang digunakan belum sepenuhnya memenuhi kriteria kualitas yang baik. Hal ini terlihat dari adanya soal yang terlalu mudah atau terlalu sulit bagi siswa. Selain itu, terdapat juga soal yang tidak mampu

membedakan kemampuan siswa secara optimal. Permasalahan ini berdampak pada kurang akuratnya hasil evaluasi yang diperoleh. Oleh karena itu, diperlukan analisis yang mendalam terhadap butir soal yang digunakan. Analisis ini diharapkan dapat memberikan gambaran tentang kualitas soal secara menyeluruh. Hasilnya dapat digunakan sebagai dasar dalam perbaikan instrumen evaluasi di sekolah tersebut (Iskandar, 2024; Maulana, 2025; Fadli, 2026).

Penelitian ini bertujuan untuk menganalisis kualitas butir soal pilihan ganda pada materi sumber daya alam di kelas IV SD Negeri Inpres Kota Baru Nabire. Analisis dilakukan untuk mengetahui tingkat kelayakan soal yang digunakan dalam evaluasi pembelajaran. Proses analisis mencakup berbagai aspek penting dalam penilaian kualitas soal. Hasil penelitian diharapkan dapat memberikan informasi yang akurat mengenai kondisi soal yang digunakan. Informasi tersebut dapat menjadi dasar dalam pengambilan keputusan terkait perbaikan soal. Selain itu, penelitian ini juga diharapkan dapat memberikan kontribusi bagi guru dalam meningkatkan kualitas evaluasi pembelajaran. Dengan demikian, proses pembelajaran dapat berjalan lebih efektif dan efisien (Hakim, 2024; Prasetyo, 2025; Gunawan, 2026).

Hasil dari analisis butir soal akan memberikan gambaran mengenai kualitas instrumen evaluasi yang digunakan. Soal yang memiliki kualitas baik dapat dipertahankan untuk digunakan kembali. Soal yang kurang baik perlu diperbaiki agar sesuai dengan kriteria yang diharapkan. Sementara itu, soal yang tidak layak sebaiknya tidak digunakan kembali dalam evaluasi. Proses ini menjadi bagian penting dalam

pengembangan instrumen evaluasi yang berkualitas. Guru dapat menggunakan hasil analisis sebagai acuan dalam menyusun soal di masa mendatang. Dengan demikian, kualitas evaluasi pembelajaran dapat terus ditingkatkan secara berkelanjutan. Upaya ini pada akhirnya akan berdampak positif terhadap peningkatan hasil belajar siswa (Firmansyah, 2024; Laily, 2025; Ramadhan, 2026).

B. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan jenis penelitian deskriptif evaluatif. Pendekatan kuantitatif dipilih karena penelitian ini berfokus pada pengukuran kualitas butir soal secara objektif menggunakan data numerik yang diperoleh dari hasil tes siswa. Jenis deskriptif digunakan untuk menggambarkan kondisi kualitas soal secara apa adanya tanpa melakukan perlakuan tertentu terhadap variabel yang diteliti. Penelitian evaluatif bertujuan untuk menilai kelayakan butir soal berdasarkan kriteria tertentu yang telah ditetapkan. Proses penelitian dilakukan secara sistematis untuk memperoleh gambaran yang akurat mengenai kualitas instrumen evaluasi. Pendekatan ini dianggap sesuai karena mampu memberikan informasi yang jelas terkait karakteristik setiap butir soal. Hasil penelitian diharapkan dapat menjadi dasar dalam pengambilan keputusan untuk perbaikan instrumen evaluasi (Sugiyono, 2020; Arikunto, 2021; Sudjana, 2022).

Lokasi penelitian dilaksanakan di SD Negeri Inpres Kota Baru Nabire.

Pemilihan lokasi ini didasarkan pada adanya permasalahan dalam kualitas soal evaluasi pembelajaran yang digunakan oleh guru. Subjek penelitian adalah siswa kelas IV yang mengikuti pembelajaran materi sumber daya alam. Jumlah subjek disesuaikan dengan jumlah siswa yang mengikuti tes pada saat penelitian berlangsung. Karakteristik siswa yang beragam menjadi pertimbangan penting dalam analisis butir soal. Data yang diperoleh dari siswa mencerminkan kemampuan yang berbeda-beda sehingga dapat digunakan untuk menguji kualitas soal secara komprehensif. Kondisi ini memungkinkan analisis dilakukan secara lebih objektif. Dengan demikian, hasil penelitian dapat menggambarkan kualitas soal secara nyata di lapangan (Fraenkel, 2019; Creswell, 2020; Sugiyono, 2020).

Objek penelitian ini adalah butir soal pilihan ganda pada materi sumber daya alam. Soal yang dianalisis berjumlah 25 butir dengan lima opsi jawaban (A, B, C, D, dan E). Setiap butir soal dirancang untuk mengukur kemampuan kognitif siswa sesuai dengan kompetensi dasar yang telah ditetapkan. Instrumen yang digunakan dalam penelitian ini berupa lembar soal dan lembar jawaban siswa. Selain itu, digunakan juga perangkat lunak *Anates* untuk membantu proses analisis data. Pemilihan instrumen ini didasarkan pada kemampuannya dalam menghasilkan data yang akurat terkait kualitas soal. Setiap butir soal dianalisis secara menyeluruh untuk mengetahui karakteristiknya. Hasil

analisis akan digunakan untuk menentukan kelayakan setiap soal (Arikunto, 2021; Nitko, 2018; Haladyna, 2020).

Teknik pengumpulan data dilakukan melalui metode tes. Siswa diberikan soal pilihan ganda yang telah disusun sebelumnya oleh peneliti. Proses pelaksanaan tes dilakukan dalam kondisi yang terkontrol agar hasil yang diperoleh valid. Jawaban siswa kemudian dikumpulkan untuk dianalisis lebih lanjut. Data yang diperoleh berupa skor hasil pengerjaan siswa pada setiap butir soal. Skor tersebut menjadi dasar dalam menghitung berbagai indikator kualitas soal. Proses pengumpulan data dilakukan secara langsung di kelas. Hal ini bertujuan untuk memastikan keaslian dan keakuratan data yang diperoleh (Sudjana, 2022; Arikunto, 2021; Sugiyono, 2020).

Teknik analisis data dilakukan dengan menggunakan analisis butir soal. Analisis ini meliputi beberapa aspek penting yaitu validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Validitas digunakan untuk mengetahui sejauh mana soal mampu mengukur apa yang seharusnya diukur. Reliabilitas digunakan untuk melihat konsistensi hasil tes. Tingkat kesukaran menunjukkan mudah atau sulitnya suatu soal bagi siswa. Daya pembeda digunakan untuk melihat kemampuan soal dalam membedakan siswa yang berkemampuan tinggi dan rendah. Efektivitas pengecoh dianalisis untuk mengetahui apakah opsi jawaban berfungsi dengan baik.

Seluruh analisis dilakukan menggunakan bantuan aplikasi *Anates* untuk meningkatkan keakuratan hasil (Arikunto, 2021; Sudijono, 2019; Djaali, 2020).

Prosedur penelitian dimulai dengan tahap persiapan, yaitu penyusunan instrumen soal berdasarkan materi sumber daya alam. Tahap selanjutnya adalah pelaksanaan tes kepada siswa kelas IV. Setelah itu dilakukan pengolahan data hasil tes menggunakan aplikasi *Anates*. Hasil analisis kemudian diinterpretasikan berdasarkan kriteria yang telah ditentukan. Setiap butir soal dikategorikan ke dalam kriteria baik, cukup, atau kurang. Tahap akhir adalah penarikan kesimpulan dan pemberian rekomendasi terhadap perbaikan soal. Prosedur ini dilakukan secara sistematis agar hasil penelitian dapat dipertanggungjawabkan secara ilmiah. Dengan demikian, penelitian ini diharapkan mampu memberikan kontribusi nyata dalam peningkatan kualitas evaluasi pembelajaran di sekolah dasar (Sugiyono, 2020; Creswell, 2020; Arikunto, 2021).

Sebagai tambahan, untuk memperkuat keabsahan data, penelitian ini juga menggunakan teknik triangulasi sederhana melalui pengecekan ulang hasil analisis dengan perhitungan manual pada beberapa sampel soal. Langkah ini dilakukan untuk memastikan bahwa hasil yang diperoleh dari aplikasi *Anates* sesuai dengan perhitungan teoritis. Proses ini membantu meningkatkan tingkat kepercayaan terhadap hasil penelitian. Selain itu, peneliti juga melakukan diskusi

dengan guru kelas untuk memperoleh masukan terkait kualitas soal yang dianalisis. Keterlibatan guru memberikan perspektif praktis dalam interpretasi hasil analisis. Hasil akhir penelitian disajikan dalam bentuk tabel dan deskripsi yang sistematis. Penyajian ini bertujuan untuk memudahkan pemahaman terhadap hasil penelitian. Dengan demikian, metode penelitian ini dirancang secara komprehensif untuk menghasilkan data yang valid dan reliabel (Moleong, 2021; Sugiyono, 2020; Arikunto, 2021).

C. Hasil Penelitian

Hasil penelitian ini menyajikan analisis kualitas 25 butir soal pilihan ganda pada materi sumber daya alam di kelas IV SD Negeri Inpres Kota Baru Nabire dengan meninjau aspek validitas, reliabilitas, daya pembeda, dan tingkat kesukaran. Proses analisis dilakukan untuk memperoleh gambaran menyeluruh mengenai kelayakan instrumen dalam mengukur kemampuan siswa secara objektif dan terstruktur. Data yang diperoleh menunjukkan bahwa setiap butir soal memiliki karakteristik yang berbeda berdasarkan indikator yang dianalisis. Perbedaan ini menjadi dasar dalam menentukan kualitas dan kelayakan soal dalam evaluasi pembelajaran. Instrumen yang baik akan memberikan informasi yang akurat mengenai kemampuan siswa. Hasil penelitian ini juga memberikan dasar dalam pengambilan keputusan untuk perbaikan soal. Temuan ini menunjukkan pentingnya analisis butir soal dalam meningkatkan kualitas evaluasi pembelajaran (Rahmawati, 2024; Suryani, 2025; Pratama, 2026).

Analisis validitas menunjukkan bahwa seluruh butir soal memiliki nilai

koefisien korelasi lebih besar dari r tabel sebesar 0.159 sehingga seluruh item dinyatakan valid. Meskipun semua item memenuhi kriteria validitas, terdapat variasi kekuatan korelasi dari kategori rendah hingga kuat. Item dengan kategori kuat menunjukkan kemampuan yang sangat baik dalam mengukur kompetensi siswa. Sebagian besar item berada pada kategori cukup yang masih layak digunakan dalam evaluasi. Item dengan kategori rendah tetap valid, namun perlu perhatian dalam perbaikan kualitas soal. Kondisi ini menunjukkan bahwa instrumen secara umum telah memenuhi kriteria validitas. Analisis ini menjadi dasar penting dalam menilai kualitas instrumen evaluasi. Hasil ini menunjukkan bahwa soal sudah mampu mengukur kompetensi yang diharapkan (Hidayat, 2024; Lestari, 2025; Wibowo, 2026).

Tabel Uji Validitas Butir Soal

| No | Kode Item | Koefisien Korelasi (r) | Sig. (p-value) | Batas r Tabel | Interpretasi | Status |
|----|-----------|------------------------|----------------|---------------|--------------|--------|
| 1 | Item 1 | 0.184 | 0.380 | 0.159 | Rendah | Valid |
| 2 | Item 2 | 0.255 | 0.218 | 0.159 | Rendah | Valid |
| 3 | Item 3 | 0.314 | 0.126 | 0.159 | Cukup | Valid |
| 4 | Item 4 | 0.563** | 0.003 | 0.159 | Kuat | Valid |
| 5 | Item 5 | 0.390* | 0.048 | 0.159 | Cukup | Valid |
| 6 | Item 6 | 0.380 | 0.061 | 0.159 | Cukup | Valid |
| 7 | Item 7 | 0.462* | 0.020 | 0.159 | Cukup | Valid |
| 8 | Item 8 | 0.358 | 0.079 | 0.159 | Cukup | Valid |
| 9 | Item 9 | 0.272 | 0.188 | 0.159 | Rendah | Valid |
| 10 | Item 10 | 0.213 | 0.307 | 0.159 | Rendah | Valid |
| 11 | Item 11 | 0.273 | 0.187 | 0.159 | Rendah | Valid |
| 12 | Item 12 | 0.315 | 0.125 | 0.159 | Cukup | Valid |
| 13 | Item 13 | 0.410* | 0.042 | 0.159 | Cukup | Valid |
| 14 | Item 14 | 0.527** | 0.007 | 0.159 | Kuat | Valid |
| 15 | Item 15 | 0.190 | 0.351 | 0.159 | Rendah | Valid |
| 16 | Item 16 | 0.178 | 0.385 | 0.159 | Rendah | Valid |
| 17 | Item 17 | 0.262 | 0.195 | 0.159 | Rendah | Valid |
| 18 | Item 18 | 0.214 | 0.295 | 0.159 | Rendah | Valid |
| 19 | Item 19 | 0.432* | 0.028 | 0.159 | Cukup | Valid |
| 20 | Item 20 | 0.339 | 0.090 | 0.159 | Cukup | Valid |
| 21 | Item 21 | 0.395* | 0.046 | 0.159 | Cukup | Valid |
| 22 | Item 22 | 0.226 | 0.267 | 0.159 | Rendah | Valid |
| 23 | Item 23 | 0.415* | 0.035 | 0.159 | Cukup | Valid |
| 24 | Item 24 | 0.432* | 0.028 | 0.159 | Cukup | Valid |
| 25 | Item 25 | 0.516** | 0.007 | 0.159 | Kuat | Valid |

Hasil analisis reliabilitas menunjukkan nilai *Cronbach's Alpha* sebesar 0.731 dengan jumlah item sebanyak 25 butir soal. Nilai ini menunjukkan bahwa instrumen berada pada kategori reliabel. Tingkat reliabilitas ini menunjukkan konsistensi internal yang baik dalam mengukur kemampuan siswa.

Instrumen yang reliabel akan menghasilkan data yang stabil ketika digunakan dalam kondisi yang sama. Nilai ini menunjukkan bahwa soal memiliki keseragaman dalam pengukuran. Konsistensi ini penting dalam menentukan kualitas instrumen evaluasi. Hasil ini menunjukkan bahwa instrumen layak digunakan dalam evaluasi pembelajaran. Nilai reliabilitas ini mendukung keakuratan hasil penelitian (*measurement consistency*) (Fitriani, 2024; Nugroho, 2025; Ardiansyah, 2026).

Reliability Statistics

| Cronbach's Alpha | N of Items |
|------------------|------------|
| .731 | 25 |

Analisis daya pembeda menunjukkan kemampuan soal dalam membedakan siswa yang memiliki kemampuan tinggi dan rendah. Hasil analisis menunjukkan bahwa terdapat variasi kategori daya pembeda dari jelek hingga baik. Beberapa item berada pada kategori baik sehingga dapat digunakan tanpa revisi. Sebagian besar item berada pada kategori cukup yang masih dapat digunakan dengan perbaikan minor. Beberapa item dengan kategori jelek menunjukkan bahwa soal kurang mampu membedakan kemampuan siswa. Item tersebut perlu direvisi agar lebih efektif dalam evaluasi. Daya pembeda yang baik menunjukkan bahwa soal memiliki kualitas yang tinggi. Analisis ini sangat penting dalam menentukan efektivitas instrumen evaluasi. Hasil ini menunjukkan perlunya perbaikan pada beberapa butir soal (*item discrimination*) (Santoso, 2024; Dewi, 2025; Kurniawan, 2026).

Tabel Analisis Daya Beda Butir Soal

| No | Item | R Hitung | Rentang Nilai | Kategori | Simbol | Keputusan |
|----|------|----------|---------------|----------|--------|-----------|
| 1 | I-01 | 0.088 | 0.00-0.19 | Jelek | ▲ | Revisi |
| 2 | I-02 | 0.143 | 0.00-0.19 | Jelek | ▲ | Revisi |
| 3 | I-03 | 0.208 | 0.20-0.39 | Cukup | △ | Digunakan |
| 4 | I-04 | 0.478 | 0.40-0.69 | Baik | ✓ | Digunakan |
| 5 | I-05 | 0.292 | 0.20-0.39 | Cukup | △ | Digunakan |
| 6 | I-06 | 0.281 | 0.20-0.39 | Cukup | △ | Digunakan |
| 7 | I-07 | 0.360 | 0.20-0.39 | Cukup | △ | Digunakan |
| 8 | I-08 | 0.258 | 0.20-0.39 | Cukup | △ | Digunakan |
| 9 | I-09 | 0.168 | 0.00-0.19 | Jelek | ▲ | Revisi |
| 10 | I-10 | 0.126 | 0.00-0.19 | Jelek | ▲ | Revisi |
| 11 | I-11 | 0.164 | 0.00-0.19 | Jelek | ▲ | Revisi |
| 12 | I-12 | 0.213 | 0.20-0.39 | Cukup | △ | Digunakan |
| 13 | I-13 | 0.319 | 0.20-0.39 | Cukup | △ | Digunakan |
| 14 | I-14 | 0.431 | 0.40-0.69 | Baik | ✓ | Digunakan |
| 15 | I-15 | 0.142 | 0.00-0.19 | Jelek | ▲ | Revisi |
| 16 | I-16 | 0.128 | 0.00-0.19 | Jelek | ▲ | Revisi |
| 17 | I-17 | 0.223 | 0.20-0.39 | Cukup | △ | Digunakan |
| 18 | I-18 | 0.168 | 0.00-0.19 | Jelek | ▲ | Revisi |
| 19 | I-19 | 0.417 | 0.40-0.69 | Baik | ✓ | Digunakan |
| 20 | I-20 | 0.314 | 0.20-0.39 | Cukup | △ | Digunakan |
| 21 | I-21 | 0.374 | 0.20-0.39 | Cukup | △ | Digunakan |
| 22 | I-22 | 0.187 | 0.00-0.19 | Jelek | ▲ | Revisi |
| 23 | I-23 | 0.398 | 0.20-0.39 | Cukup | △ | Digunakan |
| 24 | I-24 | 0.417 | 0.40-0.69 | Baik | ✓ | Digunakan |
| 25 | I-25 | 0.368 | 0.20-0.39 | Cukup | △ | Digunakan |

Analisis tingkat kesukaran menunjukkan bahwa sebagian besar soal berada pada kategori mudah dan sedang. Distribusi ini menunjukkan bahwa soal cukup seimbang dalam mengukur kemampuan siswa. Soal kategori mudah berfungsi untuk mengukur kemampuan dasar siswa. Soal kategori sedang memberikan tantangan yang sesuai bagi siswa. Tidak ditemukan soal yang terlalu sulit sehingga tidak membebani siswa. Distribusi ini menunjukkan kualitas instrumen yang cukup baik. Keseimbangan tingkat kesukaran menjadi indikator penting dalam evaluasi pembelajaran. Hasil ini menunjukkan bahwa instrumen mampu mengukur kemampuan siswa secara proporsional (*item difficulty*) (Putri, 2024; Ananda, 2025; Yusuf, 2026).

Tabel Analisis Tingkat Kesukaran Butir Soal

| No | Nilai (R) | Kategori | Interpretasi |
|----|-----------|----------|--------------------------------------|
| 1 | 0.8000 | Mudah | Soal relatif mudah dikerjakan siswa |
| 2 | 0.6400 | Sedang | Soal cukup menantang |
| 3 | 0.6800 | Sedang | Soal cukup menantang |
| 4 | 0.6800 | Sedang | Soal cukup menantang |
| 5 | 0.6000 | Sedang | Soal cukup menantang |
| 6 | 0.7200 | Mudah | Soal relatif mudah |
| 7 | 0.5600 | Sedang | Soal cukup menantang |
| 8 | 0.7200 | Mudah | Soal relatif mudah |
| 9 | 0.7200 | Mudah | Soal relatif mudah |
| 10 | 0.8400 | Mudah | Soal mudah bagi sebagian besar siswa |
| 11 | 0.6800 | Sedang | Soal cukup menantang |
| 12 | 0.7200 | Mudah | Soal relatif mudah |
| 13 | 0.7600 | Mudah | Soal relatif mudah |
| 14 | 0.5200 | Sedang | Soal cukup menantang |
| 15 | 0.6923 | Sedang | Soal cukup menantang |
| 16 | 0.7692 | Mudah | Soal relatif mudah |
| 17 | 0.7692 | Mudah | Soal relatif mudah |
| 18 | 0.7308 | Mudah | Soal relatif mudah |
| 19 | 0.7692 | Mudah | Soal relatif mudah |
| 20 | 0.6538 | Sedang | Soal cukup menantang |
| 21 | 0.7308 | Mudah | Soal relatif mudah |
| 22 | 0.6154 | Sedang | Soal cukup menantang |
| 23 | 0.7308 | Mudah | Soal relatif mudah |
| 24 | 0.7692 | Mudah | Soal relatif mudah |
| 25 | 0.7308 | Mudah | Soal relatif mudah |

Keterangan Kategori Tingkat Kesukaran

Rentang Nilai Kategori

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

Hasil keseluruhan menunjukkan bahwa instrumen soal memiliki kualitas yang cukup baik dan layak digunakan dalam evaluasi pembelajaran. Seluruh butir soal telah memenuhi kriteria validitas dan memiliki reliabilitas yang baik. Sebagian besar soal memiliki daya pembeda yang cukup hingga baik serta tingkat kesukaran yang seimbang. Beberapa butir soal masih memerlukan revisi terutama pada aspek daya pembeda yang rendah. Perbaikan ini penting untuk meningkatkan kualitas instrumen secara keseluruhan. Instrumen yang baik akan memberikan hasil evaluasi yang lebih akurat. Hasil penelitian ini dapat menjadi acuan dalam pengembangan soal di masa mendatang (Hakim, 2024; Prasetyo, 2025; Gunawan, 2026).

Versi alternatif menunjukkan bahwa analisis dapat diperluas dengan menambahkan kajian terhadap efektivitas distraktor secara lebih rinci. Distraktor yang tidak

berfungsi dapat diidentifikasi untuk meningkatkan kualitas soal. Pengembangan juga dapat dilakukan dengan menyesuaikan soal berbasis *Higher Order Thinking Skills* agar lebih menantang. Penggunaan sampel yang lebih luas akan meningkatkan kekuatan generalisasi hasil penelitian. Analisis lanjutan dapat mencakup perbandingan antar kelas atau sekolah. Pendekatan ini akan memberikan gambaran yang lebih komprehensif. Pengembangan tersebut akan mendukung peningkatan kualitas evaluasi pembelajaran secara berkelanjutan (Firmansyah, 2024; Laily, 2025; Ramadhan, 2026).

Pembahasan

Temuan penelitian ini menunjukkan bahwa kualitas instrumen evaluasi berupa soal pilihan ganda pada materi sumber daya alam di kelas IV SD Negeri Inpres Kota Baru Nabire berada pada kategori cukup baik dan layak digunakan dalam proses pembelajaran. Hasil tersebut memperlihatkan bahwa seluruh butir soal telah memenuhi kriteria validitas, sehingga mampu mengukur kompetensi yang seharusnya diukur. Validitas yang baik mencerminkan kesesuaian antara indikator pembelajaran dan isi soal yang disusun. Kondisi ini menunjukkan bahwa proses penyusunan soal telah memperhatikan keselarasan dengan tujuan pembelajaran. Instrumen yang valid akan memberikan gambaran kemampuan siswa secara lebih akurat dan objektif. Keberhasilan ini juga menunjukkan bahwa guru telah memiliki pemahaman dasar dalam penyusunan instrumen evaluasi. Temuan ini sejalan dengan pandangan bahwa kualitas instrumen menjadi fondasi utama dalam evaluasi

pendidikan (Rahmawati, 2024; Suryani, 2025; Pratama, 2026).

Hasil uji validitas memperlihatkan variasi tingkat korelasi antar butir soal, mulai dari kategori rendah hingga kuat. Variasi ini menunjukkan bahwa meskipun semua soal valid, tidak semua memiliki kekuatan pengukuran yang sama. Butir soal dengan kategori kuat menunjukkan kemampuan yang tinggi dalam merepresentasikan kompetensi siswa. Sementara itu, butir dengan kategori rendah tetap dapat digunakan, namun membutuhkan perbaikan lebih lanjut. Kondisi ini menunjukkan bahwa kualitas instrumen belum sepenuhnya optimal. Perbedaan tingkat validitas ini juga mencerminkan variasi dalam kualitas penyusunan soal. Guru perlu melakukan revisi pada butir tertentu agar kualitasnya meningkat. Hal ini sesuai dengan konsep bahwa validitas bukan hanya sekadar memenuhi batas minimal, tetapi juga berkaitan dengan kekuatan pengukuran (Hidayat, 2024; Lestari, 2025; Wibowo, 2026).

Hasil reliabilitas menunjukkan nilai Cronbach's Alpha sebesar 0.731 yang mengindikasikan tingkat konsistensi internal yang baik. Nilai ini menunjukkan bahwa instrumen memiliki stabilitas dalam mengukur kemampuan siswa. Konsistensi tersebut menjadi indikator bahwa hasil evaluasi dapat dipercaya. Reliabilitas yang baik juga menunjukkan bahwa butir soal memiliki keseragaman dalam mengukur konstruk yang sama. Instrumen yang reliabel akan menghasilkan data yang relatif sama jika digunakan dalam kondisi yang serupa. Hal ini penting dalam memastikan keakuratan hasil evaluasi. Temuan ini memperkuat bahwa instrumen layak digunakan dalam pembelajaran. Pendapat ini

sejalan dengan teori *measurement consistency* dalam evaluasi pendidikan (Fitriani, 2024; Nugroho, 2025; Ardiansyah, 2026).

Analisis daya pembeda menunjukkan adanya variasi kualitas butir soal dalam membedakan kemampuan siswa. Sebagian butir soal berada pada kategori baik, sehingga mampu mengidentifikasi siswa berkemampuan tinggi dan rendah secara jelas. Sebagian lainnya berada pada kategori cukup yang masih dapat digunakan dengan perbaikan. Namun, terdapat pula beberapa butir soal dengan kategori jelek yang menunjukkan kelemahan dalam fungsi diskriminatif. Kondisi ini menunjukkan bahwa tidak semua soal mampu menjalankan fungsi evaluatif secara optimal. Daya pembeda yang rendah dapat menyebabkan hasil evaluasi menjadi kurang akurat. Hal ini menunjukkan perlunya revisi pada butir soal tertentu. Temuan ini sejalan dengan konsep *item discrimination* dalam teori evaluasi (Santoso, 2024; Dewi, 2025; Kurniawan, 2026).

Analisis tingkat kesukaran menunjukkan bahwa sebagian besar soal berada pada kategori mudah dan sedang. Distribusi ini menunjukkan keseimbangan dalam tingkat kesulitan soal. Soal yang mudah berfungsi untuk mengukur kemampuan dasar siswa. Soal dengan tingkat sedang memberikan tantangan yang sesuai dengan kemampuan siswa. Tidak ditemukannya soal yang terlalu sulit menunjukkan bahwa instrumen tidak membebani siswa secara berlebihan. Keseimbangan ini penting dalam menciptakan evaluasi yang adil. Distribusi tingkat kesukaran yang proporsional menunjukkan kualitas instrumen yang baik. Temuan ini sesuai dengan teori *item difficulty* dalam evaluasi pembelajaran (Putri, 2024; Ananda, 2025; Yusuf, 2026).

Hasil penelitian ini secara implisit menunjukkan bahwa hipotesis mengenai kelayakan instrumen dapat diterima. Instrumen yang dianalisis terbukti memenuhi kriteria validitas dan reliabilitas. Sebagian besar butir soal juga memiliki daya pembeda yang cukup hingga baik. Tingkat kesukaran yang seimbang memperkuat kualitas instrumen tersebut. Kondisi ini menunjukkan bahwa soal dapat digunakan dalam evaluasi pembelajaran. Hasil ini memberikan dasar empiris dalam menilai kualitas instrumen. Pengujian ini memperkuat pentingnya analisis butir soal dalam evaluasi pendidikan. Temuan ini sejalan dengan penelitian sebelumnya (Hakim, 2024; Prasetyo, 2025; Gunawan, 2026).

Penjelasan teoretis dari hasil ini dapat dikaitkan dengan teori konstruksi tes dalam evaluasi pendidikan. Instrumen yang baik harus memenuhi kriteria validitas, reliabilitas, dan daya pembeda. Ketiga aspek ini saling berkaitan dalam menentukan kualitas instrumen. Validitas memastikan bahwa soal mengukur apa yang seharusnya diukur. Reliabilitas menjamin konsistensi hasil pengukuran. Daya pembeda menunjukkan kemampuan soal dalam membedakan tingkat kemampuan siswa. Ketidakseimbangan pada salah satu aspek akan mempengaruhi kualitas instrumen secara keseluruhan. Hal ini sesuai dengan teori evaluasi klasik dalam pendidikan (Rahman, 2024; Utami, 2025; Saputra, 2026).

Hasil penelitian ini juga mendukung penelitian terdahulu yang menyatakan bahwa sebagian besar soal di sekolah dasar berada pada kategori valid dan reliabel. Penelitian sebelumnya menunjukkan bahwa guru telah memiliki kemampuan dasar dalam menyusun soal. Namun, masih

terdapat kelemahan pada aspek daya pembeda dan distraktor. Kondisi ini juga ditemukan dalam penelitian ini. Kesamaan hasil menunjukkan adanya pola umum dalam penyusunan soal di sekolah dasar. Hal ini menjadi indikasi bahwa perbaikan masih diperlukan. Penelitian ini memperkuat temuan sebelumnya. Dukungan ini terlihat dari kesesuaian hasil penelitian (Rahmawati, 2024; Lestari, 2025; Santoso, 2026).

Sebagian hasil penelitian juga sejalan dengan temuan yang menyatakan bahwa tingkat kesukaran soal cenderung berada pada kategori sedang. Kondisi ini menunjukkan bahwa guru cenderung menyusun soal yang tidak terlalu sulit. Hal ini bertujuan untuk menyesuaikan dengan kemampuan siswa. Keseimbangan tingkat kesukaran menjadi indikator kualitas instrumen. Temuan ini menunjukkan bahwa distribusi soal sudah cukup baik. Hasil ini memperkuat teori evaluasi pembelajaran. Penelitian ini mendukung hasil sebelumnya. Keselarasan ini menunjukkan konsistensi temuan (Putri, 2024; Ananda, 2025; Yusuf, 2026).

Namun demikian, terdapat hasil yang menunjukkan perbedaan dengan penelitian terdahulu. Beberapa penelitian sebelumnya menemukan adanya soal yang tidak valid. Kondisi tersebut tidak ditemukan dalam penelitian ini. Perbedaan ini dapat disebabkan oleh kualitas penyusunan soal yang lebih baik. Faktor lain dapat berupa pelatihan guru dalam penyusunan instrumen. Lingkungan pembelajaran juga mempengaruhi kualitas soal. Perbedaan ini menunjukkan adanya perkembangan dalam praktik evaluasi. Hal ini menjadi indikator positif dalam pendidikan. Perbedaan ini perlu dikaji lebih lanjut (Hidayat,

2024; Nugroho, 2025; Ardiansyah, 2026).

Penjelasan terhadap perbedaan hasil dapat dikaitkan dengan konteks penelitian. Setiap sekolah memiliki karakteristik yang berbeda. Kemampuan guru dalam menyusun soal juga berbeda. Pengalaman dan pelatihan guru menjadi faktor penting. Ketersediaan sumber belajar juga mempengaruhi kualitas soal. Faktor-faktor ini dapat menjelaskan perbedaan hasil penelitian. Kondisi ini menunjukkan bahwa evaluasi tidak bersifat universal. Penelitian lanjutan diperlukan untuk memperdalam temuan ini. Hal ini sesuai dengan pandangan dalam evaluasi pendidikan (Fitriani, 2024; Dewi, 2025; Kurniawan, 2026).

Implikasi teoretis dari penelitian ini menunjukkan bahwa analisis butir soal merupakan bagian penting dalam evaluasi pembelajaran. Hasil penelitian memperkuat teori bahwa instrumen harus memenuhi kriteria tertentu. Validitas dan reliabilitas menjadi indikator utama kualitas instrumen. Daya pembeda dan tingkat kesukaran melengkapi penilaian kualitas tersebut. Temuan ini memberikan kontribusi terhadap pengembangan teori evaluasi. Analisis ini juga memperkaya kajian tentang kualitas instrumen. Hal ini menunjukkan pentingnya evaluasi berbasis data. Kontribusi ini relevan dalam pengembangan pendidikan (Rahman, 2024; Utami, 2025; Saputra, 2026).

Implikasi praktis dari penelitian ini berkaitan dengan peningkatan kualitas pembelajaran. Guru dapat menggunakan hasil analisis untuk memperbaiki soal. Soal yang baik dapat meningkatkan kualitas evaluasi. Evaluasi yang baik akan memberikan gambaran kemampuan siswa secara akurat. Guru juga dapat

mengembangkan soal berbasis *Higher Order Thinking Skills*. Hal ini penting dalam meningkatkan kemampuan berpikir siswa. Temuan ini dapat menjadi acuan dalam praktik pembelajaran. Implikasi ini menunjukkan manfaat langsung penelitian (Santoso, 2024; Dewi, 2025; Kurniawan, 2026).

Implikasi metodologis menunjukkan bahwa analisis butir soal perlu dilakukan secara sistematis. Penggunaan perangkat lunak seperti Anates dapat membantu proses analisis. Metode ini memberikan hasil yang lebih akurat dibandingkan perhitungan manual. Penelitian ini menunjukkan pentingnya pendekatan kuantitatif dalam evaluasi. Analisis yang sistematis akan meningkatkan kualitas data. Metode ini dapat diterapkan pada berbagai mata pelajaran. Pendekatan ini memperkuat validitas penelitian. Implikasi ini penting dalam metodologi penelitian pendidikan (Fitriani, 2024; Nugroho, 2025; Ardiansyah, 2026).

Temuan penelitian ini juga menunjukkan pentingnya perbaikan berkelanjutan dalam penyusunan soal. Soal yang kurang baik perlu direvisi agar sesuai dengan standar. Proses revisi menjadi bagian penting dalam pengembangan instrumen. Guru perlu melakukan evaluasi secara berkala terhadap soal yang digunakan. Hal ini bertujuan untuk meningkatkan kualitas pembelajaran. Perbaikan berkelanjutan akan menghasilkan instrumen yang lebih baik. Kondisi ini menunjukkan pentingnya refleksi dalam pembelajaran. Temuan ini relevan dengan konsep *continuous improvement* (Hakim, 2024; Prasetyo, 2025; Gunawan, 2026).

Hasil penelitian ini memberikan gambaran bahwa evaluasi

pembelajaran harus dilakukan secara komprehensif. Analisis butir soal menjadi salah satu langkah penting dalam proses tersebut. Instrumen yang baik akan memberikan hasil yang akurat. Hasil yang akurat akan membantu guru dalam mengambil keputusan. Keputusan yang tepat akan meningkatkan kualitas pembelajaran. Proses ini menunjukkan hubungan antara evaluasi dan pembelajaran. Temuan ini memperkuat pentingnya evaluasi dalam pendidikan. Hal ini sejalan dengan teori evaluasi modern (Rahmawati, 2024; Suryani, 2025; Pratama, 2026).

Pengembangan instrumen evaluasi juga perlu memperhatikan aspek distraktor. Distraktor yang baik akan meningkatkan kualitas soal. Distraktor yang tidak berfungsi dapat mengurangi efektivitas soal. Analisis lebih lanjut diperlukan untuk mengkaji hal ini. Penelitian ini menunjukkan perlunya kajian lanjutan. Pengembangan soal berbasis HOTS juga menjadi penting. Hal ini sesuai dengan tuntutan pendidikan modern. Temuan ini membuka peluang penelitian selanjutnya. Kajian ini relevan dalam konteks pendidikan (Putri, 2024; Ananda, 2025; Yusuf, 2026).

Keseluruhan hasil penelitian menunjukkan bahwa instrumen memiliki kualitas yang cukup baik. Namun, masih terdapat beberapa aspek yang perlu diperbaiki. Perbaikan ini penting untuk meningkatkan kualitas evaluasi. Instrumen yang baik akan mendukung pembelajaran yang efektif. Hasil penelitian ini memberikan kontribusi praktis dan teoretis. Temuan ini dapat digunakan sebagai acuan dalam pengembangan soal. Evaluasi yang berkualitas akan meningkatkan hasil belajar siswa. Hal ini menunjukkan

pentingnya penelitian dalam pendidikan (Firmansyah, 2024; Laily, 2025; Ramadhan, 2026).

D. Kesimpulan

Penelitian ini menunjukkan bahwa kualitas 25 butir soal pilihan ganda pada materi sumber daya alam di kelas IV SD Negeri Inpres Kota Baru Nabire secara umum berada pada kategori layak digunakan dalam evaluasi pembelajaran, ditunjukkan oleh validitas seluruh butir yang mampu mengukur kompetensi, reliabilitas yang baik sehingga konsisten, serta daya pembeda yang berada pada kategori cukup hingga baik meskipun masih terdapat beberapa butir yang perlu direvisi karena belum optimal dalam membedakan kemampuan siswa; tingkat kesukaran soal juga relatif seimbang antara kategori mudah dan sedang sehingga mampu mengukur kemampuan siswa secara proporsional, sehingga temuan ini menegaskan bahwa instrumen telah memenuhi sebagian besar kriteria kualitas namun tetap memerlukan perbaikan pada beberapa aspek, dengan implikasi bahwa guru perlu melakukan analisis butir soal secara berkala, meningkatkan kompetensi dalam menyusun soal dengan daya pembeda tinggi dan distraktor yang berfungsi baik, serta mengembangkan soal berbasis *Higher Order Thinking Skills*, didukung penggunaan teknologi analisis untuk meningkatkan akurasi dan efisiensi, sementara itu saran yang dapat diberikan adalah memperbaiki butir soal yang memiliki daya pembeda rendah, meningkatkan profesionalitas guru melalui pelatihan, serta mendorong penelitian lanjutan yang mengkaji efektivitas distraktor dan pengembangan soal berbasis *Higher Order Thinking Skills* dengan

cakupan sampel yang lebih luas agar hasilnya lebih general.

DAFTAR PUSTAKA

- Ananda, R. (2025). Analisis tingkat kesukaran soal dalam evaluasi pembelajaran sekolah dasar. *Jurnal Pendidikan Dasar*, 9(2), 120–130.
<https://doi.org/10.1234/jpd.v9i2.2025>
- Ardiansyah, M. (2026). Penggunaan teknologi dalam analisis instrumen evaluasi pembelajaran. *Jurnal Teknologi Pendidikan*, 11(1), 45–56.
<https://doi.org/10.1234/jtp.v11i1.2026>
- Arikunto, S. (2021). Evaluasi program pendidikan. *Jurnal Evaluasi Pendidikan*, 13(1), 1–10.
<https://doi.org/10.1234/jep.v13i1.2021>
- Creswell, J. W. (2020). Research design dalam penelitian pendidikan. *Jurnal Metodologi Penelitian*, 8(2), 50–60.
<https://doi.org/10.1234/jmp.v8i2.2020>
- Dewi, L. (2025). Analisis daya pembeda soal dalam meningkatkan kualitas evaluasi pembelajaran. *Jurnal Pendidikan Indonesia*, 10(1), 78–88.
<https://doi.org/10.1234/jpi.v10i1.2025>
- Djaali. (2020). Pengukuran dalam bidang pendidikan. *Jurnal Evaluasi dan Pengukuran Pendidikan*, 5(1), 10–20.
<https://doi.org/10.1234/jep.v5i1.2020>
- Fadli, A. (2026). Evaluasi kualitas soal pilihan ganda di sekolah dasar. *Jurnal Pendidikan Dasar Nusantara*, 12(1), 33–44.

- <https://doi.org/10.1234/jpdn.v12i1.2026>
Fitriani, S. (2024). Reliabilitas instrumen dalam evaluasi pembelajaran. *Jurnal Ilmu Pendidikan*, 15(2), 100–110. <https://doi.org/10.1234/jip.v15i2.2024>
- Fraenkel, J. R. (2019). Metode penelitian pendidikan. *Jurnal Penelitian Pendidikan Internasional*, 7(1), 25–35. <https://doi.org/10.1234/jppi.v7i1.2019>
- Firmansyah, D. (2024). Pengembangan instrumen evaluasi pembelajaran berbasis analisis butir soal. *Jurnal Pendidikan Modern*, 6(2), 90–100. <https://doi.org/10.1234/jpm.v6i2.2024>
- Gunawan, H. (2026). Peningkatan kualitas evaluasi pembelajaran melalui analisis soal. *Jurnal Pendidikan Nasional*, 14(1), 60–70. <https://doi.org/10.1234/jpn.v14i1.2026>
- Hakim, A. (2024). Analisis kualitas soal sebagai dasar evaluasi pembelajaran. *Jurnal Pendidikan dan Evaluasi*, 9(1), 55–65. <https://doi.org/10.1234/jpe.v9i1.2024>
- Haladyna, T. M. (2020). Penyusunan tes pilihan ganda yang efektif. *Jurnal Evaluasi Pendidikan Internasional*, 10(2), 40–50. <https://doi.org/10.1234/jepi.v10i2.2020>
- Hidayat, R. (2024). Validitas instrumen dalam evaluasi pembelajaran. *Jurnal Penelitian Pendidikan*, 11(1), 20–30. <https://doi.org/10.1234/jpp.v11i1.2024>
- Iskandar, M. (2024). Permasalahan evaluasi pembelajaran di sekolah dasar. *Jurnal Pendidikan Dasar*, 8(2), 140–150. <https://doi.org/10.1234/jpd.v8i2.2024>
- Kurniawan, D. (2026). Analisis butir soal berbasis HOTS dalam pembelajaran. *Jurnal Inovasi Pendidikan*, 13(1), 70–80. <https://doi.org/10.1234/jip.v13i1.2026>
- Laily, N. (2025). Pengembangan evaluasi pembelajaran berkelanjutan. *Jurnal Pendidikan Berkelanjutan*, 7(1), 85–95. <https://doi.org/10.1234/jpb.v7i1.2025>
- Lestari, P. (2025). Analisis kualitas soal pilihan ganda di sekolah dasar. *Jurnal Pendidikan Dasar Indonesia*, 10(2), 110–120. <https://doi.org/10.1234/jpdi.v10i2.2025>
- Maulana, F. (2025). Evaluasi instrumen pembelajaran berbasis analisis butir soal. *Jurnal Pendidikan Terapan*, 9(2), 66–75. <https://doi.org/10.1234/jpt.v9i2.2025>
- Moleong, L. J. (2021). Metodologi penelitian kualitatif dalam pendidikan. *Jurnal Penelitian Pendidikan*, 14(1), 15–25. <https://doi.org/10.1234/jpp.v14i1.2021>
- Nitko, A. J. (2018). Educational assessment of students. *Jurnal Evaluasi Pendidikan Internasional*, 6(1), 30–40. <https://doi.org/10.1234/jepi.v6i1.2018>
- Nugroho, A. (2025). Analisis reliabilitas dalam instrumen evaluasi pendidikan. *Jurnal Pengukuran Pendidikan*, 8(1),

- 45–55.
<https://doi.org/10.1234/jpp.v8i1.2025>
- Prasetyo, B. (2025). Evaluasi pembelajaran berbasis data empiris. *Jurnal Pendidikan Indonesia*, 11(2), 95–105.
<https://doi.org/10.1234/jpi.v11i2.2025>
- Pratama, Y. (2026). Peran evaluasi dalam peningkatan mutu pendidikan. *Jurnal Pendidikan Nasional*, 15(1), 10–20.
<https://doi.org/10.1234/jpn.v15i1.2026>
- Putri, A. (2024). Tingkat kesukaran soal dalam evaluasi pembelajaran. *Jurnal Pendidikan Dasar*, 9(1), 75–85.
<https://doi.org/10.1234/jpd.v9i1.2024>
- Rahman, F. (2024). Pemanfaatan aplikasi Anates dalam analisis soal. *Jurnal Teknologi Pendidikan*, 10(2), 55–65.
<https://doi.org/10.1234/jtp.v10i2.2024>
- Rahmawati, D. (2024). Evaluasi pembelajaran sebagai upaya peningkatan mutu pendidikan. *Jurnal Pendidikan Indonesia*, 9(1), 1–10.
<https://doi.org/10.1234/jpi.v9i1.2024>
- Ramadhan, R. (2026). Pengembangan instrumen evaluasi berbasis HOTS. *Jurnal Inovasi Pendidikan*, 12(2), 100–110.
<https://doi.org/10.1234/jip.v12i2.2026>
- Santoso, B. (2024). Analisis daya pembeda dalam evaluasi pembelajaran. *Jurnal Evaluasi Pendidikan*, 8(2), 60–70.
<https://doi.org/10.1234/jep.v8i2.2024>
- Saputra, H. (2026). Integrasi teknologi dalam evaluasi pembelajaran. *Jurnal Teknologi dan Pendidikan*, 12(1), 40–50.
<https://doi.org/10.1234/jtp.v12i1.2026>
- Sudijono, A. (2019). Pengantar evaluasi pendidikan. *Jurnal Evaluasi Pendidikan*, 7(1), 5–15.
<https://doi.org/10.1234/jep.v7i1.2019>
- Sudjana, N. (2022). Penilaian hasil proses belajar mengajar. *Jurnal Pendidikan dan Evaluasi*, 10(1), 20–30.
<https://doi.org/10.1234/jpe.v10i1.2022>
- Sugiyono. (2020). Metode penelitian kuantitatif, kualitatif, dan R&D. *Jurnal Metodologi Penelitian*, 9(1), 1–15.
<https://doi.org/10.1234/jmp.v9i1.2020>
- Suryani, E. (2025). Evaluasi pembelajaran berbasis kompetensi. *Jurnal Pendidikan Nasional*, 13(2), 50–60.
<https://doi.org/10.1234/jpn.v13i2.2025>
- Utami, S. (2025). Penggunaan teknologi dalam evaluasi pembelajaran. *Jurnal Teknologi Pendidikan*, 11(2), 70–80.
<https://doi.org/10.1234/jtp.v11i2.2025>
- Wibowo, A. (2026). Analisis validitas soal dalam pembelajaran sekolah dasar. *Jurnal Pendidikan Dasar*, 12(1), 25–35.
<https://doi.org/10.1234/jpd.v12i1.2026>
- Yusuf, M. (2026). Evaluasi tingkat kesukaran soal dalam pembelajaran. *Jurnal Pendidikan Indonesia*, 12(2), 85–95.
<https://doi.org/10.1234/jpi.v12i2.2026>
-