

## Analisis Validitas dan Reliabilitas Soal IPA Berbasis Literasi Sains dan Pemahaman Konsep Sistem Pencernaan Manusia

Muhammad Yusuf<sup>1</sup>, Nova E. Ntobuo<sup>2</sup>, Nur Azizah Dj. Junus<sup>3\*</sup>, Nyoman Sriastusti<sup>4</sup>, Risna H. Panyo<sup>5</sup>

<sup>12345</sup>Departement of Natural Science, Universitas Negeri Gorontalo, Gorontalo, Indonesia

Email: [1muhammad.yusuf@ung.ac.id](mailto:1muhammad.yusuf@ung.ac.id), [2novantobuo@ung.ac.id](mailto:2novantobuo@ung.ac.id),

[3nur\\_s2ilmupengetahuanalam@mahasiswa.ung.ac.id](mailto:3nur_s2ilmupengetahuanalam@mahasiswa.ung.ac.id),

[4nyoman\\_s2ilmupengetahuanalam@mahasiswa.ung.ac.id](mailto:4nyoman_s2ilmupengetahuanalam@mahasiswa.ung.ac.id),

[5risna\\_s2ilmupengetahuanalam@mahasiswa.ung.ac.id](mailto:5risna_s2ilmupengetahuanalam@mahasiswa.ung.ac.id)

Received: 9 April 2026

Revised: 14 April 2026

Accepted: 24 April 2026

### Abstract

This study aims to analyze the quality of science test items based on scientific literacy and conceptual understanding of the human digestive system material seen from the validity, reliability, level of difficulty, and discriminatory power. This study uses a descriptive quantitative approach. Data analysis used descriptive statistics with the help of SPSS. Validity test uses Product Moment correlation at a significance level of 5%, while reliability uses the Cronbach Alpha coefficient. The level of difficulty is seen from the P index, and discriminatory power is used to assess the ability to differentiate between high and low ability participants. The results showed that the number of valid items was 55% with a reliability of 0.864 (very high), the level of difficulty of 90% was in the moderate category, and 80% of discriminatory power had good to very good discriminatory power. Overall, this instrument meets the criteria of good quality, and can be used as an evaluation tool to measure students' conceptual understanding and scientific literacy consistently and accurately.

**Keywords:** *Validity, Reliability, Scientific Literacy, Concept Understanding*

### Abstrak

Penelitian ini bertujuan untuk menganalisis kualitas butir soal IPA berbasis literasi sains dan pemahaman konsep materi sistem pencernaan manusia dilihat dari validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Penelitian ini menggunakan pendekatan kuantitatif deskriptif. Analisis data yang digunakan statistik deskriptif dengan bantuan SPSS. Uji validitas menggunakan korelasi *Product Moment* pada taraf signifikansi 5%, sedangkan reliabilitas menggunakan koefisien *Alpha Cronbach*. Tingkat kesukaran dilihat dari indeks P, dan daya pembeda digunakan untuk menilai kemampuan dalam membedakan peserta berkemampuan tinggi dan rendah. Hasil penelitian menunjukkan jumlah butir yang valid adalah 55% dengan reliabilitas 0,864 (sangat tinggi), tingkat kesukaran 90% berada pada kategori sedang, dan daya pembeda 80% memiliki daya pembeda baik sampai sangat baik. Secara keseluruhan, instrumen ini memenuhi kriteria kualitas yang baik, dan dapat dijadikan sebagai sarana evaluasi guna mengukur pemahaman konsep dan literasi sains siswa dengan konsisten dan akurat.

**Kata kunci:** Validitas, Reliabilitas, Literasi Sains, Pemahaman Konsep

## PENDAHULUAN

Literasi sains dan konsep yang sangat kuat dalam siswa harus diangkut dalam pendidikan abad ke-21, sehingga siswa dapat beradaptatif dengan kemajuan ilmu dan teknologi yang tak terbayangkan cepat. Seperti yang dijelaskan oleh Wilsa et al., (2023) sains sebenarnya berarti lebih daripada sekedar mengetahui apa yang terdapat di dalam suatu bahan. Daripada itu, literasi sains juga menandakan kemampuan manusia untuk menentukan argumen ilmiah yang punya substansi yang memadai, menilai apakah suatu sumber informasi bisa diandalkan atau tidak, dan menentukan langkah yang harus diambil

berdasarkan data dan bukti secara kritis dan rasional. Kemampuan ini kunci jadi karenanya salah satu indikator utama yang diukur oleh penilaian kompetensi sains internasional seperti PISA. Sayangnya, siswa Indonesia telah menunjukkan tingkat literasi sains yang relatif rendah, dan kesimpang ini juga dibenarkan dengan fakta bahwa instrumen evaluasi belum sepenuhnya mempertimbangkan dimensi literasi (Dewi, 2022).

Selanjutnya, pemahaman konsep merupakan dasar dari proses belajar mengajar sains. Ketika siswa memiliki pemahaman konsep yang kuat, mereka dapat mengaitkan pengetahuan dengan fenomena di kehidupan sehari-hari dan menggunakannya secara kontekstual (Vonny et al., 2021). Namun, sebagian besar siswa umumnya kurang dapat paham terhadap konsep yang diajarkan, seperti sistem pencernaan manusia secara menyeluruh. Ini sebagian besar karena bahan ajar dan media pembelajaran yang ada kurang mendukung literasi sains dan pemahaman konsep (Elu et al., 2024). Karena itu, instrumen evaluasi yang sesuai dan selaras untuk literasi sains dan pemahaman konsep sangat diperlukan.

Instrumen evaluasi yang berkualitas ditandai dengan terpenuhinya aspek validitas dan reliabilitas, agar instrumen dapat memberikan informasi yang tepat terkait kemampuan siswa (Bano, 2022). Hal ini karena validitas menentukan seberapa besar instrumen mampu memperoleh pengalokasian konstruk yang dimaksudkan, sementara reliabilitas merujuk pada konsistensi hasil dari pengukuran itu sendiri. Instrumen yang berbasis literasi sains secara konsisten memberikan kualitas crossover measurement dan quality of mock-up dan sampel sesuai dengan tes subyektivitas (Dewi, 2022; Wilsa et al., 2023).

Dari uraian tersebut, selanjutnya diperlukan analisis validitas dan reliabilitas pada soal IPA berbasis literasi sains dan pemahaman konsep pada materi sistem pencernaan manusia. Penelitian ini akan memberikan gambaran mengenai kualitas butir soal tersebut guna mempergunakan alat ukur tersebut sebagai instrumen merepresentasi agar menjadi alat pengukur kemampuan literasi sains dan pemahaman konsep siswa yang valid dan reliabel. Sehingga, diharapkan penelitian ini menghasilkan informasi yang jelas dan terstruktur kepada guru untuk menyusun soal evaluasi yang bisa mempergunakan instrumen tersebut untuk melakukan pengukuran identifikasi tidak tinggi pada skala kognitif mahir atau tinggi tetapi bisa merangsang kemampuan berpikir berdasarkan tuntutan abad 21.

## **METODE PENELITIAN**

Penelitian ini menggunakan pendekatan analisis instrumen dalam metode kuantitatif deskriptif untuk mendeskripsikan kualitas soal IPA berbasis literasi sains dan pemahaman konsep pada materi sistem pencernaan manusia. Desain penelitian evaluatif terhadap 20 butir soal pilihan ganda yang dirancang mengacu pada kisi-kisi indikator kognitif, pemahaman konsep, literasi sains. Adapun kisi-kisi indikator soal dapat dilihat pada Tabel 1. Gambaran kualitas instrumen dalam mengukur literasi sains, disajikan satu contoh soal yang merepresentasikan kemampuan analisis (C4) siswa.

<b>Nomor</b>	<b>Soal</b>	<b>Kunci Jawaban</b>
13	Di dalam mulut, nasi yang dikunyah terasa sedikit manis setelah beberapa saat. Fenomena ini terjadi karena ... a. Air liur melarutkan nasi menjadi gula sederhana b. Enzim amilase mengubah amilum menjadi maltosa c. Air liur mengandung asam yang memecah protein d. Amilase bekerja memecah lemak menjadi gliserol	B

**Tabel 1.** Kisi-Kisi Instrumen Soal Berdasarkan Ranah Kognitif, Pemahaman Konsep dan Literasi Sains

Ranah Kognitif	Indikator Soal	Pemahaman Konsep	Literasi Sains	No Soal
C1 (Mengingat)	Menyebutkan organ pencernaan	Mengingat konsep dasar	Mengenali fakta ilmiah	1,2,3
C2 (Memahami)	Menjelaskan fungsi organ	Memahami konsep	Interpretasi ilmiah	4,5,6,7,8
C3 (Menerapkan)	Menentukan jalur pencernaan	Aplikasi konsep	Aplikasi konteks	9,10,11,12
C4 (Menganalisis)	Analisis enzim dan zat makanan	Relasi konsep	Penalaran ilmiah	13,14,15,16
C5 (Mengevaluasi)	Evaluasi kebiasaan makan	Penilaian konsep	Pengambilan keputusan	17,18
C6 (Mencipta)	Merancang solusi kesehatan pencernaan	Integrasi konsep	Kreativitas sains	19,20

Selanjutnya, prosedur penelitian meliputi penyusunan, uji coba, pengumpulan skor, dan analisis kualitas soal. Subjek penelitian adalah 18 siswa kelas VIII dengan skor dikotomi, yaitu 1 untuk respons yang benar, 0 untuk respons yang salah. Data dianalisis secara deskriptif menggunakan SPSS 27, dengan uji validitas Product Moment, reliabilitas Alpha Cronbach/KR-20, tingkat kesukaran, dan daya pembeda. Hasilnya sebagai landasan dalam menilai mutu soal sebagai instrument penilaian pembelajaran IPA.

## HASIL DAN PEMBAHASAN

### Uji Validitas

Validitas isi merupakan tahap awal penyusunan instrumen yang menunjukkan kesesuaian butir soal dengan indikator kompetensi. Kisi-kisi telah mengintegrasikan pemahaman konsep dan literasi sains sehingga seluruh indikator terwakili. Secara teoretis, literasi sains mencakup kemampuan memahami masalah, menguraikan fenomena, dan menggunakan bukti ilmiah, sejalan dengan Azizah & Widowati Budijastuti, (2022) bahwa penilaian harus mengukur analisis dan interpretasi. Selain itu, tanpa validitas isi, pengukuran dapat menjadi bias dan tidak akurat (Prasetya & Pratama, 2023).

Berdasarkan hasil perhitungan uji validitas pada 20 butir soal pilihan ganda pada materi Sistem Pencernaan Manusia, diperoleh bahwa 11 butir soal tergolong valid dan 9 butir soal tidak valid, sebagaimana ditunjukkan pada Gambar 1.

Berdasarkan hasil uji validitas dari instrumen soal pada Tabel 1 diperoleh  $r_{tabel} = 0,497$  ( $\alpha = 0,05$ ;  $n = 18$ ). Dikatakan butir valid jika  $r_{hitung} > 0,497$ . Berdasarkan hasil perhitungan, butir yang dianggap valid sebanyak 11 butir (55%) tepatnya pada soal nomor 2, 5, 7, 9, 10, 13, 14, 15, 16, 18, dan 20 dan 9 buah soal tidak valid. Penyebab tidak validnya soal mungkin karena indikator yang tidak sesuai, redaksi kurang valid, atau distraktor yang tidak optimal.



**Gambar 1.** Persentase (%) Uji Validitas Butir Soal Materi Sistem Pencernaan Manusia

Soal yang valid lebih akurat dalam mencerminkan penguasaan konsep siswa, namun soal yang tidak valid memungkinkan terjadi bias penilaian. Maka seharusnya soal yang tidak valid direvisi atau dihilangkan. Soal yang valid dapat dihimpun ke dalam kumpulan soal sebagai arsip, tetapi soal yang tidak valid sebaiknya dihapuskan dan direntangkan kembali menjadi soal yang valid (Anazar & Munir, 2024; Dianova & Anwar, 2024; Rachmawati & Pradana, 2025).

Dengan demikian sebagian besar butir yang disajikan telah valid dan dapat digunakan untuk mengukur pemahaman konsep lengkap dan literasi ilmiah pada pokok bahasan sistem pencernaan manusia. Namun, butir yang belum valid perlu digunakan kembali dan diperbaiki.

### Uji Reliabilitas

Reliabilitas menunjukkan konsistensi skor pengukuran, yaitu pengukuran ulang pada kondisi sejenis, maka hasil kembali cenderung sama. Tarmizi et al., (2021) menyatakan bahwa reliabilitas berkaitan dengan akurasi dan konsistensi skor tes. Hasil analisis 20 butir soal menunjukkan 11 butir valid dan 9 tidak valid, dengan uji reliabilitas pada 11 butir valid menghasilkan nilai 0,864, sebagaimana di tunjukkan pada Tabel 2.

**Tabel 2.** Hasil Uji Reliabilitas pada Butir yang Valid

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
<b>0.864</b>	<b>0.865</b>	<b>11</b>

Hasil pengukuran reliabilitas menggunakan SPSS menunjukkan nilai *Cronbach's Alpha* sebesar 0,864 pada 11 butir soal valid. Nilai ini termasuk kategori sangat tinggi (0,80–1,00) dan memenuhi kriteria reliabel karena  $> 0,70$ . Hasil ini sesuai dengan penelitian Anazar & Munir, (2024); Pratama et al., (2023) sehingga menunjukkan konsistensi internal instrumen yang sangat baik.

Reliabilitas yang tinggi menunjukkan bahwa instrumen dapat memberikan data yang konsisten dan akurat dalam pengukuran terhadap pemahaman konsep dan literasi sains mahasiswa, menjadi penentu dalam proses pengambilan keputusan pembelajaran yang lebih lanjutan, seperti evaluasi strategi mengajar dan memetakan remediasi penguasaan penguasaan materi mahasiswa. Oleh sebab itu, Uji reliabilitas dapat dijadikan prasyarat utama agar data tes ini dapat dijadikan alat ukur yang stabil dan bermakna, sebuah pengetahuan yang tersedia dalam beberapa penelitian (Dianova & Anwar, 2024;

Rachmawati & Pradana, 2025). Oleh karena itu, instrumen yang dikembangkan telah memenuhi standar kelayakannya sebagai alat evaluasi pembelajaran dalam literasi pada materi sistem pencernaan manusia.

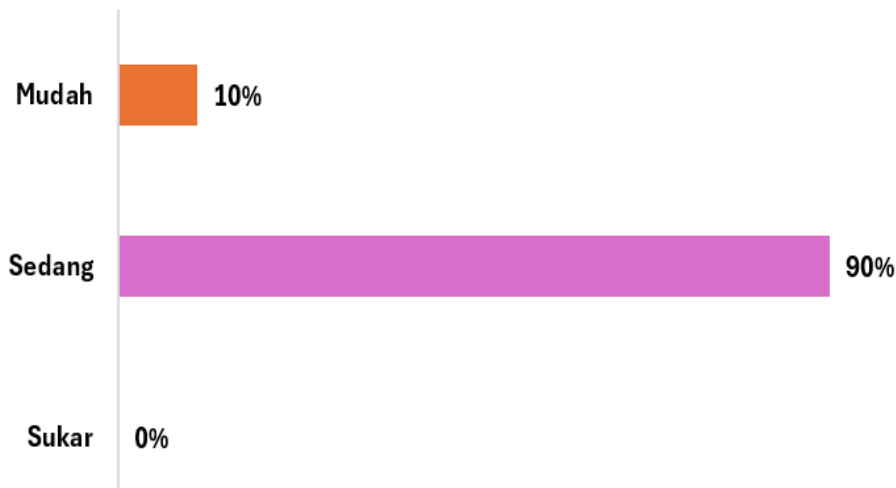
### Tingkat Kesukaran Soal

Tingkat kesukaran mencerminkan tingkat kesulitan atau kemudahan pertanyaan tersebut dan mencakup sifat pertanyaan (Sarassanti & Wahidah, 2024). Tingkat kesulitan pertanyaan terkait adalah proporsi kuesioner yang berhasil menjawab pertanyaan dengan benar. Semakin rendah nilai indeks kesukaran yang dihasilkan dari perhitungan, maka dapat dikatakan soal tersebut lebih sukar (Dianova & Anwar, 2024). Adapun kriteria penafsiran tingkat kesukaran soal dapat dilihat pada **Tabel 3**. (Sarassanti & Wahidah, 2024).

**Tabel 3.** Interpretasi Tingkat Kesukaran Butir Soal

Interval	Interpretasi
0.00 – 0.30	Soal Sukar
0.31 – 0.70	Soal Sedang
0.71 – 1.00	Soal Mudah

Berdasarkan hasil uji tingkat kesukaran pada 20 butir soal materi *Sistem Pencernaan Manusia*, diperoleh indeks kesukaran (P) yang bervariasi antara 0,339 hingga 0,778. Sebagaimana ditunjukkan pada **Gambar 2**.



**Gambar 2.** Persentase (%) Tingkat Kesukaran Materi Sistem Pencernaan Manusia

Berdasarkan 20 butir soal, terdapat 10% soal mudah dan 90% soal tingkat menengah tanpa soal sulit. Tingkat kesulitan secara keseluruhan masih baik karena semuanya masih dalam kisaran ideal umum yaitu skor 0,30-0,70 (Anazar & Munir, 2024; Faridah, 2021).

Butir soal yang dikategorikan sebagai tingkat menengah dianggap efektif dalam mengukur kemampuan siswa dan layak disimpan dalam bank soal. Namun, distribusinya dianggap tidak ideal karena tidak ada soal dengan tingkat kesulitan tinggi, yang merupakan alokasi 3-4-3 atau 3-5-2 terhadap rasio ideal (Sarassanti & Wahidah, 2024).

Oleh karena itu, instrumen ini perlu ditambahkan lebih banyak soal dengan tingkat kesulitan tinggi untuk menyeimbangkan serta mengukur semua aspek kemampuan siswa secara lebih menyeluruh.

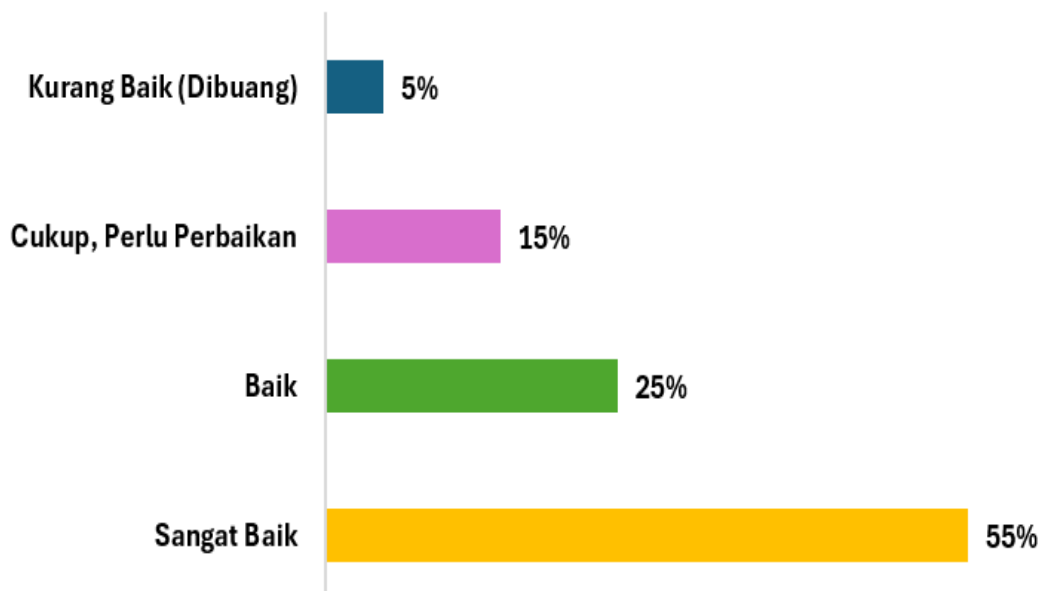
### Daya Pembeda

Daya pembeda soal menunjukkan sejauh mana suatu butir soal mampu mengelompokkan siswa berdasarkan tingkat pemahaman materi, yaitu antara yang memahami dan yang belum memahami. Menurut Tarmizi et al., (2021) daya pembeda penting untuk menilai apakah sebuah soal optimal dalam mengukur kemampuan siswa secara akurat. Semakin besar koefisien daya pembeda, semakin memiliki daya guna suatu soal untuk mengidentifikasi perbedaan kemampuan siswa. Penilaian koefisien daya pembeda mengacu pada kriteria yang tercantum dalam **Tabel 4.** (Imam & Syahrial, 2023; Sarassanti & Wahidah, 2024).

**Tabel 4.** Kriteria Tingkat Daya Pembeda Soal.

Kriteria Daya Pembeda	
0.40 atau lebih	Sangat Baik
0.30 – 0.39	Baik
0.20 – 0.29	Cukup, Perlu Perbaikan
0.19 – ke bawah	Kurang Baik (Dibuang)

Adapun hasil uji daya pembeda terhadap 20 butir soal materi sistem pencernaan manusia, diperoleh indeks daya pembeda (D) yang bervariasi antara 0,222 hingga 0,778 sebagaimana ditunjukkan pada **Gambar 3.**



**Gambar 3.** Persentase Daya Pembeda Soal pada Materi Sistem Pencernaan Manusia

Pengujian daya pembeda mengindikasikan bahwa dari 20 butir soal, 55% berkategori sangat baik, 25% baik, 15% cukup atau harus ada perbaikan dan 5% kurang baik. Soal dengan daya pembeda baik-sangat baik dapat efektif membedakan kemampuan siswa pada rentang 0,41–1,00 serta tetap berkualitas jika di atas 0,70 (Marambaawang et al., 2023; Pohan & Nurrahmah, 2024; Salsabila et al., 2025).

Sementara hanya kategori cukup perlu disempurnakan dan yang kurang baik lebih harus dibuang karena indikator tidak terdefinisi dengan jelas agar kesukarannya proporsional (Nurhalimah et al., 2022). Karena itu, Maka dari itu, mayoritas butir soal itulah sebabnya mayoritas butir soal memiliki daya pembeda yang baik hingga sangat baik sehingga instrumen telah menepati kualitas, meskipun butir-butir tertentu perlu diperbaiki agar hasil pengukuran lebih maksimal.

## SIMPULAN

Dari hasil analisis, seumum, dari 20 butir soal IPA materi sistem pencernaan manusia instrumen sudah berada pada kategori soal yang berkualitas baik. Butir soal valid sebesar 55% dan untuk butir yang butuh revisi (tidak valid) 45%, reliabilitasnya sangat tinggi cronbach's alpha 0,864. Kesukuran dominan dari kategori sedang 90% dan mudah 10%, soal butuh revisi, dan soal sukar belum ada. Daya pembeda banyak yang tergolong pada kategori baik hingga sangat baik. Hal ini menandakan bahwa instrumen tersebut cukup efektif untuk dimanfaatkan sebagai perangkat penilaian, namun butuh penyempurnaan pada butir-butir tertentu dan penambahan butir soal sukar.

## DAFTAR PUSTAKA

- Anazar, & Munir, M. (2024). Analisis Kualitas Butir Soal Ujian Akhir Semester Genap Mata Pelajaran Simulasi Dan Komunikasi Digital Kelas X Multi Media di SMK Negeri 4 Kendari Tahun Ajaran 2021/2022. *Journal Of Information Technology And Education (Jited)*, 2(1), 41–51.
- Azizah, Z. N., & Widowati Budijastuti. (2022). Pengembangan Instrumen Penilaian Untuk Mengukur Keterampilan Literasi. *Bioedu Berkala Ilmiah Pendidikan Biologi*, 11(1), 89–97.
- Bano, V. O. (2022). Volume : 8 Bulan : Februari Tahun : 2022 Analisis Kriteria Butir Soal Ujian Sekolah Mata Pelajaran IPA di SMP Negeri 1 Waingapu Volume : 8 Nomor : 1 Bulan : Februari Tahun : 2022. *Jurnal Ideas: Pendidikan, Sosial Dan Budaya*, 8, 145–152. <https://doi.org/10.32884/ideas.v8i1.660>
- Dewi, K. (2022). Pengembangan Instrumen Tes Untuk Mengukur Kemampuan Literasi Sains Siswa Kelas V Sekolah Dasar Pada Materi Sistem Pencernaan Manusia. *Jurnal Kualita Pendidikan*, 3(3), 140–145.
- Dianova, F. R., & Anwar, N. (2024). Analisis Butir Uji Validitas, Reliabilitas, Tingkat Kesukaran, Dan Daya Pembeda Soal Sumatif Bahasa Arab Sd Islam. *Jurnal Bahasa Daerah Indonesia*, 3, 1–13. <https://doi.org/10.47134/jbdi.v1i3.2863>
- Elu, F. E., Ulia, V., Sila, R., & Shidik, M. A. (2024). Pengembangan Media Pembelajaran Booklet Pada Materi Sistem Pencernaan Pada Manusia Kelas VIII. *Jbioedra: Jurnal Pendidikan Biologi*, 2(8), 264–273.
- Faridah, A. (2021). Karakteristik Butir Soal Ujian Akhir Semester Mata Pelajaran Sejarah. *Jurnal Penelitian Hukum Dan Pendidikan*, 20(2), 1281–1288.
- Imam, M. K., & Syahrial. (2023). Hasil Uji Tingkat Daya Pembeda, Tingkat Kesukaran, Keberfungsian Distraktor, Validitas Dan Reliabilitas Pada Soal PKN Siswa Kelas IV di SDN 066056 Medan Denai. *Jurnal Pgsd Indonesia*, 09(Juni), 42–46.
- Khumaira, R., Ramadan, R. R., Khairunnisa, S., Jannah, M., & Marhadi, H. (2024). Analisis Daya Beda Pada Soal Tes Mata Pelajaran. 8(5), 533–540.
- Magdalena, I., Fauziah, S. N., Faziha, S. N., & Nupus, F. S. (1994). Analisis Validitas, Reliabilitas, Tingkat Kesulitan Dan Daya Beda Butir Soal Ujian Akhir Semester Tema 7 Kelas III SDN Karet 1 Sepatan. *Bintang : Jurnal Pendidikan Dan Sains*, 37(6), 452–456. <https://ejournal.stitpn.ac.id/index.php/bintang>
- Marambaawang, D. N., Bano, V. O., & Enda, R. R. H. (2023). Analisis Kualitas Butir Soal Penilaian Akhir Semester Gasal Tahun 2021/2022 Menggunakan Iteman di SMP Negeri 1 Kampera. *De\_Journal (Dharmas Education Journal)*, 4(1), 233–243.
- Nurhalimah, S., Hidayati, Y., Rosidi, I., & Hadi, W. P. (2022). The Relationship Between Item Validity With Distinguishing Power And Difficulty Level Of Multiple Choice Questions Pas (In Indonesian). *Natural Science Education Research*, 4(3), 249–257.
- Pohan, S. N. A., & Nurrahmah. (2024). Analisis Butir Soal Penilaian Akhir Semester (Pas) Bahasa Indonesia Kelas Viii Smp Negeri 18 Banda Aceh 2023/2024.
- Prasetya, W. A., & Pratama, A. T. (2023). Item Quality Analysis Using The Rasch Model To

- Measure Critical Thinking Ability In The Material Of The Human Digestive System Of Biology Subject In High School. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 27(1), 76–91.
- Pratama, F., Darvina, Y., & Dewi, W. S. (2023). Analysis Quality Of Equation Items In Learning Evaluation For Physics Subject In High Schools Of The Mentawai Island Regency. *Pillars Of Physics Education*, 16(2), 141–154.
- Rachmawati, D., & Pradana, A. B. (2025). Analisis Butir Soal Mata Pelajaran Ekonomi: Validitas, Reliabilitas, Tingkat Kesukaran, Dan Daya Pembeda. *Jurnal Pendidikan Ekonomi (Jupe)*, 13(3), 273–284. <https://doi.org/10.26740/jupe.v13n3.p273-284>
- Salsabila, A., Khairani, F., Pangestu, D., Nurwahidin, M., & Lampung, U. (2025). Pengaruh Penggunaan Bahan Ajar Liveworksheets Caxra: *Jurnal Pendidikan Sekolah Dasar*. 05(01), 339–340.
- Sarassanti, Y., & Wahidah, A. N. (2024). Analisis Butir Soal Pecahan Pada Mahasiswa: Validitas, Reliabilitas, Daya Pembeda, Tingkat Kesukaran. 175–181.
- Tarmizi, P., Setiono, P., Amaliyah, Y., & Agrian, A. (2021). Analisis Butir Soal Pilihan Ganda Tema Sehat Itu Penting Kelas V Sd Negeri 04 Kota Bengkulu. *Else (Elementary School Education Journal)*, 4, 124–132. <https://doi.org/10.30651/else.v4i2.7090>
- Vonny, Nihlah, K., Miarsyah, M., & Ristanto, R. H. (2021). Mempromosikan Literasi Biologi Kepada Siswa Sekolah Menengah: Pengembangan Instrumen Tes Untuk Kelas Vii Promoting. *Bioedusiana: Jurnal Pendidikan Biologi*, 6(2), 251–265.
- Wilsa, A. W., Rusilowati, A., Susilaningsih, E., Jaja, J., & Nurpadillah, V. (2023). Jurnal Penelitian Dan Evaluasi Pendidikan Validity , Reliability , And Item Characteristics Of Cell Material Science Literacy Assessment Instruments. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 27(2). <https://doi.org/10.21831/pep.v27i2.61577>