



Analisis Butir Soal Objektif Penilaian Sumatif Akhir Semester Gasal Mata Pelajaran Ekonomi

Nur Sitta Rahmawati¹, M. Fathur Rahman²

^{1,2}Universitas Negeri Semarang, Indonesia

E-mail: nursittarahmawati@students.unnes.ac.id, fathur@mail.unnes.ac.id

Article Info	Abstract
Article History Received: 2025-06-10 Revised: 2025-07-20 Published: 2025-08-01	<p>Item analysis is important to obtain information about the quality of items on the assessment that has been carried out. The problem at the research location is that the teacher has not conducted item analysis due to time constraints. Supported by data that shows low student learning outcomes with an average completeness below 70%. This study aims to analyze the objective question items (multiple choice) of the Gasal PSAS of Economics subject class XI in the 2024/2025 school year as many as 30 items conducted at SMA Negeri 1 Bumiayu. This research focuses on analyzing aspects of validity, reliability, discriminating power, difficulty level, and the effectiveness of distractors. The research method used descriptive quantitative. The subjects of this study were all students of class XI specialization group of economics as many as 143 students. Data analysis used the help of Anates V.4.09 software. The results of this study indicate that the objective items of the Gasal PSAS of Economics class XI SMA Negeri 1 Bumiayu in the 2024/2025 academic year have item reliability in the fairly reliable category, namely 0.52. The results of the analysis of 30 items, it is known that the validity of the questions is still low, only 23% of the questions are valid. The discriminating power of the question is 74% in the good and sufficient category. The effectiveness of distractors was 90% in the excellent, good, and good enough categories. The overall analysis concluded that this question was 67% acceptable, 20% needed to be revised, and 13% should be removed.</p>
Keywords: <i>Learning Evaluation;</i> <i>Summative Assessment;</i> <i>Item Analysis;</i> <i>Anates V.4.09.</i>	

Artikel Info	Abstrak
Sejarah Artikel Diterima: 2025-06-10 Direvisi: 2025-07-20 Dipublikasi: 2025-08-01	<p>Analisis butir soal penting dilakukan untuk mendapatkan informasi tentang kualitas butir soal pada penilaian yang sudah terlaksana. Permasalahan di lokasi penelitian adalah guru belum melakukan analisis butir soal karena keterbatasan waktu. Didukung dengan data yang menunjukkan rendahnya hasil belajar siswa dengan rata-rata ketuntasan di bawah 70%. Penelitian ini bertujuan untuk menganalisis butir soal objektif (pilihan ganda) PSAS Gasal mata pelajaran Ekonomi kelas XI tahun ajaran 2024/2025 sebanyak 30 butir yang dilakukan di SMAN 1 Bumiayu. Penelitian ini fokus pada analisis aspek validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh. Metode penelitian menggunakan deskriptif kuantitatif. Subjek penelitian ini adalah seluruh siswa kelas XI kelompok peminatan ekonomi sebanyak 143 siswa. Analisis data menggunakan bantuan software Anates V.4.09. Hasil penelitian ini menunjukkan bahwa butir soal objektif PSAS Gasal mata pelajaran Ekonomi kelas XI SMAN 1 Bumiayu Tahun Ajaran 2024/2025 memiliki reliabilitas butir soal pada kategori cukup reliabel yaitu 0,52. Hasil dari analisis 30 butir soal, diketahui bahwa validitas soal masih rendah, hanya 23% soal yang valid. Daya pembeda soal sebanyak 74% pada kategori baik dan cukup. Tingkat kesukaran soal sebanyak 56% pada kategori sedang. Efektivitas pengecoh sebanyak 90% pada kategori sangat baik, baik, dan cukup baik. Analisis secara keseluruhan menyimpulkan bahwa soal ini 67% dapat diterima, 20% perlu direvisi, dan 13% sebaiknya dibuang.</p>
Kata kunci: <i>Evaluasi Pembelajaran;</i> <i>Penilaian Sumatif;</i> <i>Analisis Butir Soal;</i> <i>Anates V.4.09.</i>	

I. PENDAHULUAN

Evaluasi pembelajaran dilakukan dengan perencanaan penilaian, pembuatan instrumen, pelaksanaan penilaian, pengolahan, pemanfaatan, dan pelaporan hasil penilaian (Fietri et al., 2021). Hasil belajar yang dievaluasi oleh guru melalui tes menjadi gambaran tentang hasil belajar yang merujuk pada standar penilaian sekolah dalam mengukur seberapa jauh program pembelajaran

yang telah dilaksanakan (Kaka et al., 2024). Setelah diadakan evaluasi hasil belajar, guru akan mendapatkan informasi dan membuat keputusan agar kualitas pembelajaran dapat meningkat. Sehingga, kegiatan evaluasi ini penting dan tidak dapat terpisahkan dari proses pendidikan, karena berkaitan dengan mutu pembelajaran dan sebagai rujukan dalam peningkatan kualitas pendidikan.

Kualitas pendidikan dapat dikatakan baik apabila seorang guru mempunyai kemampuan dalam melakukan evaluasi (Purniasari et al., 2021), sebab seorang guru tidak hanya berperan dalam membimbing, mengajar, dan mengarahkan siswa ketika belajar, tetapi juga harus melakukan evaluasi (Wulandari et al., 2025). Kemampuan mengadakan evaluasi hasil belajar menjadi salah satu bagian dari kompetensi yang perlu dimiliki oleh seorang guru kaitannya dengan kompetensi pedagogik (Farida & Musyarofah, 2021). Maka dari itu, peran guru sebagai evaluator ini sangat penting dalam upaya mencapai keberhasilan hasil belajar siswa.

Tes merupakan metode yang dapat digunakan sebagai penentuan kemampuan siswa dalam menyelesaikan tugas tertentu atau menunjukkan penguasaan siswa pada suatu keterampilan atau pengetahuan tentang pokok bahasan tertentu (Adom et al., 2020). Sebagai bagian integral dari evaluasi, tes memiliki peranan penting dalam pembelajaran (Walfath & Faurenza, 2024). Tes menjadi cara untuk mengevaluasi kompetensi siswa setelah terlibat dalam suatu pembelajaran (Umami & Rusdi, 2021). Oleh karenanya, tes digunakan dalam evaluasi untuk mengukur ketercapaian pembelajaran yang hasilnya dapat menjadi acuan dalam upaya perbaikan dan peningkatan kualitas soal.

Terdapat dua kategori tes yaitu *standardized test* dan *teacher-made test*. Tes terstandarisasi (*Standardized test*) adalah tes yang telah distandarisasi melalui pertimbangan uji validitas dan reliabilitas. Berbeda dengan tes buatan guru (*teacher-made test*) yang disusun oleh guru pengampu yang bersangkutan untuk mengetahui ketercapaian hasil belajar siswa. Tes buatan guru dirancang berdasarkan silabus dan rencana pembelajaran yang diterapkan pada suatu mata pelajaran salah satunya mata pelajaran ekonomi. Tes ini juga terbatas penggunaannya hanya pada satu sekolah. Sehingga tes buatan guru biasanya jarang melalui pertimbangan uji validitas dan reliabilitas.

Salah satu jenis tes objektif adalah soal pilihan ganda. Soal ini dapat memuat materi lebih banyak, efisien dan mempermudah guru melakukan penilaian (Nurhalimah et al., 2024). Soal pilihan ganda merupakan soal yang umumnya memiliki dua sampai lima opsi jawaban (Mania et al., 2020). Sehingga, butir soal pilihan ganda menarik untuk dianalisis.

Penilaian sumatif mengamati seberapa jauh ketercapaian tujuan pembelajaran siswa dalam kurun waktu tertentu yang mana memberikan informasi dasar tentang kesiapan siswa untuk

meneruskan ke pembelajaran tahap selanjutnya (Nirwana et al., 2024). Penilaian sumatif dalam hal ini dilakukan secara berkala setiap akhir periode pembelajaran untuk menilai secara keseluruhan hasil belajar siswa pada setiap mata pelajaran termasuk mata pelajaran ekonomi. Berikut data awal yang diperoleh dari studi pendahuluan terkait hasil belajar siswa pada PSAS Gasal kelas XI mata pelajaran Ekonomi tahun ajaran 2024/2025 di SMAN 1 Bumiayu.

Tabel 1. Ketuntasan Nilai PSAS Gasal Ekonomi

Kelas	Jumlah Siswa	Rerata Nilai	Tuntas	Tidak Tuntas
XI 7	36	54,17	17%	83%
XI 8	35	53,53	6%	94%
XI 9	36	67,06	54%	46%
XI10	36	58,69	19%	81%
Rata-Rata		58,36	24%	76%

(Sumber: Data Primer Diolah Peneliti, 2025)

Berdasarkan data di atas menunjukkan bahwa sebagian besar siswa kelas XI kelompok peminatan ekonomi memiliki ketuntasan yang rendah ditunjukkan oleh persentase yang kurang dari 70%. KKTP (Kriteria Ketuntasan Tujuan Pembelajaran) yang sekolah tetapkan adalah 70. Dari 143 siswa kelas XI kelompok peminatan ekonomi, yang memenuhi KKTP hanya berjumlah 34 siswa (24%) sedangkan sisanya sejumlah 109 siswa (76%) belum tuntas. Hal ini menandakan bahwa masih rendahnya hasil belajar pada penilaian tersebut.

Kualitas butir soal menjadi salah satu bagian penting dalam pembelajaran ekonomi, terutama di kelas XI yang sedang berada pada tahap pemahaman konsep lanjutan. Menganalisis butir soal dilaksanakan untuk diamati dan dievaluasi sejauh mana kualitas butir soal yang baik untuk dapat diaplikasikan pada tahun ajaran berikutnya (Cahyani et al., 2024). Analisis soal didefinisikan sebagai aktivitas yang mengkaji setiap butir dan item soal agar diketahui kualitas soal tersebut sehingga diperoleh soal dengan kualitas yang baik (Zalukhu et al., 2023). Kriteria yang perlu dipenuhi dalam analisis butir soal pilihan ganda diantaranya aspek validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh atau distraktor (Augustia et al., 2025).

Asmawi Nainul dan Noehi Nasution berpendapat bahwa ada empat alasan analisis butir soal diperlukan yakni karena akan 1) mengetahui kekuatan dan kelemahan butir soal; 2) menyediakan informasi tentang spesifikasi butir soal secara lengkap; 3) dapat secepatnya diketahui masalah yang terdapat dalam butir

soal; dan 4) dijadikan alat guna menilai butir soal yang akan disimpan dalam butir soal (Riani et al., 2020). Dengan demikian, informasi penting yang berkaitan dengan kualitas butir soal khususnya pada penilaian yang sudah terlaksana dapat diketahui dengan melakukan analisis butir soal.

Hasil analisis dan evaluasi soal pilihan ganda dapat lebih cepat dan praktis didapatkan dengan bantuan *software* Anates. Anates adalah salah satu perangkat lunak yang dapat digunakan untuk menganalisis butir-butir soal dengan pendekatan klasik yang ditinjau berdasarkan aspek validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh (Alista & Syahzanani, 2023). *Software* ini dikembangkan oleh seorang dosen psikologi dari Universitas Pendidikan Indonesia bernama Bapak Drs. Karno To, M.Pd bersama Bapak Yudi Wibisono, S.T. Manfaat Anates V.4.09 pada dasarnya memiliki kesamaan seperti item pengolah data lainnya yang sejenis, akan tetapi secara pengoperasiannya lebih mudah (Sari & Yudha, 2020). *Software* ini umumnya bermanfaat bagi guru dan pemerhati evaluasi pendidikan, serta bagi mahasiswa atau calon guru yang mengambil jurusan keguruan atau pendidikan.

Hasil studi pendahuluan yang telah dilakukan peneliti, diketahui bahwa SMAN 1 Bumiayu memakai soal buatan guru dalam pelaksanaan PSAS Gasal pada mata pelajaran Ekonomi kelas XI. Menyoroti temuan masalah terkait belum dilaksanakannya analisis butir soal buatan guru secara komprehensif dikarenakan terbatasnya waktu yang dimiliki guru sehingga kualitas soal tersebut masih belum diketahui. Permasalahan penelitian ini diperkuat oleh penelitian (Fiska et al., 2021) bahwa mayoritas guru membuat soal secara mendadak saat soal akan diberikan pada siswa yang berdampak perasaan terbebani ketika terlebih dahulu harus melakukan analisis butir soal karena terhambat dan terbatas pada waktu. Supaya kualitas soal khususnya pada tes objektif berupa pilihan ganda pada penilaian tersebut dapat diketahui, maka perlu adanya penelitian ini.

Perbedaan penelitian ini dengan penelitian (Saely & Shaleh, 2023) terkait penerapan alat pemrosesan data yang mana menggunakan aplikasi Winstep dan Iteman 4.3 kemudian hasilnya dibandingkan. Sementara itu, penelitian ini sejalan dengan penelitian (Alti et al., 2022) yang menggunakan *software* Anates V.4.09 sebagai alat bantu pemrosesan data. Penelitian tersebut meninjau aspek validitas, reliabilitas, daya pembeda, dan tingkat kesukaran. Namun, aspek efektivitas pengecoh belum ada. Mengacu

pada kesenjangan tersebut, menjadi kebaruan dari penelitian ini. Selain menganalisis aspek validitas, reliabilitas, daya pembeda, dan tingkat kesukaran, juga aspek efektivitas pengecoh.

Penelitian ini diharapkan dapat memberikan kontribusi pada penguatan praktik evaluasi pembelajaran melalui analisis butir soal, khususnya bagi guru dalam meningkatkan mutu soal pilihan ganda pada mata pelajaran ekonomi di tahun ajaran selanjutnya.

II. METODE PENELITIAN

1. Desain Penelitian

Penelitian ini menggunakan metode penelitian deskriptif pendekatan kuantitatif dengan jenis analisis dokumen (*documentary analysis*). Penelitian deskriptif bermaksud mengumpulkan informasi tentang keadaan suatu gejala secara apa adanya tanpa generalisasi (Hikmawati, 2020). Pemilihan pendekatan kuantitatif disebabkan data yang diperoleh dalam bentuk angka-angka yang diolah dengan *software* Anates V.4.09 dan dianalisis berdasarkan kriteria yang ditentukan kemudian ditarik kesimpulan (Riani et al., 2020).

2. Populasi dan Sampel

Populasi penelitian berjumlah 143 siswa kelas XI Peminatan Ekonomi di SMAN 1 Bumiayu. Teknik sampling yang digunakan adalah sampel jenuh, yaitu seluruh populasi dijadikan sampel karena jumlahnya relatif kecil dan dapat dijangkau sepenuhnya. Objek yang dianalisis adalah soal pilihan ganda PSAS Gasal pelajaran Ekonomi kelas XI tahun ajaran 2024/2025 sebanyak 30 soal sedangkan subjek penelitiannya adalah sebanyak 143 lembar jawaban siswa.

3. Teknik Pengumpulan Data

Pengumpulan data dilakukan melalui teknik dokumentasi, berupa lembar soal, lembar jawaban siswa, dan kunci jawaban.

4. Teknik Analisis Data

Teknik analisis data berbantuan *software* Anates versi 4.09 yang akan menghasilkan informasi berupa reliabilitas, daya pembeda, tingkat kesukaran, efektivitas pengecoh, serta korelasi skor butir dan skor total yang menunjukkan validitas.

5. Validitas

Validitas diartikan bagaimana suatu tes secara akurat dapat mengukur sesuatu yang

seharusnya diukur, jika sesuai dengan kondisi sesungguhnya maka data dianggap valid (Ikawati et al., 2024). Validitas butir soal dapat menggunakan validitas empiris dengan menghitung korelasi *point biserial*. Adapun kriterianya sebagai berikut.

$$r_{pbi} > r_{tabel} = \text{Valid}$$

$$r_{pbi} < r_{tabel} = \text{Tidak Valid}$$

6. Reliabilitas

Reliabilitas dapat diartikan sebagai keajegan/ konsistensi skor perolehan ketika dilakukan pengujian ulang pada situasi yang berbeda. Hasil perhitungan koefisien reliabilitas dapat diklasifikasikan sebagai berikut.

Sangat Reliabel	: $0,80 < r \leq 1,00$
Reliabel	: $0,60 < r \leq 0,80$
Cukup Reliabel	: $0,40 < r \leq 0,60$
Tidak Reliabel	: $0,20 < r \leq 0,40$
Sangat Tidak Reliabel	: $\leq 0,20$

Modifikasi (Ningsih et al., 2023)

7. Daya Pembeda

Daya pembeda meninjau kemampuan setiap butir soal untuk membedakan antara siswa yang mendapat skor tinggi (unggul) dengan siswa yang mendapat skor rendah (asor). Mengacu pada (Mania et al., 2020), nilai dari indeks daya pembeda (Dp) yang diperoleh dapat dikelompokkan menjadi kategori sangat baik ($0,70 < Dp \leq 1,00$), baik ($0,40 < Dp \leq 0,70$), cukup ($0,20 < Dp \leq 0,40$), jelek ($0,00 < Dp \leq 0,20$), dan sangat jelek ($Dp \leq 0,00$).

8. Tingkat Kesukaran

Tingkat kesukaran butir soal diartikan sebagai perhitungan proporsi antara siswa yang bisa menjawab benar suatu butir soal dengan jumlah seluruh siswa yang mengerjakan tes. Mengacu pada (Zalukhu et al., 2023), nilai dari indeks tingkat kesukaran (P) yang diperoleh dapat dikelompokkan menjadi kategori mudah ($P > 0,70$), sedang ($0,30 \leq P \leq 0,70$), dan sukar ($P < 0,30$).

9. Efektivitas Pengecoh

Efektivitas pengecoh bertujuan menganalisis pola suatu jawaban dengan menghitung jumlah siswa yang memilih tiap opsi jawaban di setiap butir soal. Mengacu pada modifikasi (Karindi & Rofi'i, 2024), hasil dari analisis tersebut dapat dikelompokkan menjadi kategori sangat baik (4 pengecoh

berfungsi), baik (3 pengecoh berfungsi), cukup (2 pengecoh berfungsi), kurang baik (1 pengecoh berfungsi), dan tidak baik (tidak ada pengecoh yang berfungsi).

III. HASIL DAN PEMBAHASAN

A. Hasil Penelitian

1. Validitas

Hasil penelitian terkait validitas butir soal PSAS Gasal Ekonomi menggunakan *software* Anates V.4.09, tertera pada tabel di bawah ini.

Tabel 1. Distribusi Validitas Soal PSAS

Kategori Soal	Nomor Soal	Persentase (Jumlah Soal)
Valid ($\geq 0,349$)	9,10,16,17,20,28, 29	23% (7 butir)
Tidak Valid ($< 0,349$)	1,2,3,4,5,6,7,8,11,12, 13,14,15,18, 19,21,22,23,24, 25,26,27,30	77% (23 butir)

Sumber: Data Diolah Peneliti, 2025

Hasil analisis validitas butir soal pada *software* Anates V.4.09, dapat dilihat pada hasil koefisien korelasi skor butir dengan skor total. Valid atau tidaknya butir soal dibedakan menjadi dua kategori yaitu signifikan atau sangat signifikan yang berarti soal valid dan bertanda (-) yang berarti tidak signifikan atau soal tidak valid. Nilai r_{tabel} dapat diketahui pada keterangan batas signifikansi koefisien korelasi. Sesuai kriteria, butir soal dikatakan valid ketika $r_{hitung} \geq r_{tabel}$. Berdasarkan hasil pengolahan data dengan Anates V.4.09, validitas butir soal dinyatakan baik apabila $r_{hitung} \geq 0,349$, sesuai batas signifikansi dengan $n = 30$ pada taraf 5%. Sebaliknya, nilai di bawah kriteria tersebut menandakan soal tidak valid.

Distribusi hasil validitas dari 30 butir soal objektif PSAS Gasal Ekonomi, terdapat 7 butir soal (23%) yang berada pada kategori valid, yaitu pada butir soal nomor 9, 10, 16, 17, 20, 28, dan 29 dengan nilai koefisien korelasi atau r_{hitung} pada masing-masing butir yaitu 0,419; 0,374; 0,417; 0,464; 0,349; 0,421; 0,492; sementara 23 butir soal (77%) dinyatakan tidak valid dengan skor $r_{hitung} < 0,349$. Korelasi untuk butir soal tidak valid berada di rentang (-0,145 – 0,347).

Rendahnya proporsi soal valid (23%) mengindikasikan bahwa mayoritas butir soal tidak mampu mengukur konstruk secara memadai. Hal ini menunjukkan perlunya revisi substansial dalam penyusunan soal PSAS Gasal Ekonomi untuk tahun ajaran berikutnya.

2. Reliabilitas

Hasil penelitian terkait nilai reliabilitas butir soal objektif PSAS Gasal Ekonomi menggunakan *software* Anates V.4.09, tertera pada tabel berikut.

Reliabilitas	0,52
Rata-Rata	13,48
Simpangan Baku	3,43
Korelasi XY	0,35

Sumber: Anates V.4.09

Berdasarkan aspek reliabilitas tes, hasil analisis butir soal objektif PSAS Gasal Ekonomi kelas XI menggunakan *software* Anates V.4.09 menunjukkan nilai 0,52 berada dalam rentang $0,40 < r \leq 0,60$. Nilai reliabilitas sebesar 0,52 menunjukkan bahwa konsistensi antarbutir soal tergolong "cukup reliabel", namun masih belum memenuhi kategori "reliabel tinggi". Hal ini perlu menjadi perhatian dalam penyusunan soal untuk memastikan stabilitas hasil penilaian siswa.

3. Daya Pembeda

Hasil analisis butir soal objektif PSAS Gasal Ekonomi menggunakan *software* Anates V.4.09 pada aspek daya pembeda dalam penelitian ini dibedakan menjadi 5 kategori yaitu sangat baik, baik, cukup, jelek, dan sangat jelek. Hasil penelitian terkait daya pembeda soal objektif PSAS Gasal Ekonomi menggunakan *software* Anates V.4.09, tertera pada tabel berikut.

Tabel 2. Distribusi Daya Pembeda Soal PSAS

Kategori Soal	Nomor Soal	Persentase (Jumlah Soal)
Baik ($0,40 < Dp \leq 0,70$)	9,10,16,17,20, 21,22,27,28, 29	34% (10 butir)
Cukup ($0,20 < Dp \leq 0,40$)	1,2,3,4,5,6,7,13, 18,23,24,30	40% (12 butir)
Jelek ($0,00 < Dp \leq 0,20$)	11,12,14,26	13% (4 butir)
Sangat Jelek ($Dp \leq 0,00$)	8,15,19,25	13% (4 butir)

Sumber: Data Diolah Peneliti, 2025

Hasil analisis daya pembeda butir soal objektif PSAS Gasal Ekonomi kelas XI SMAN 1 Bumiayu berdasarkan pada kriteria yang ada menunjukkan bahwa dari 30 butir soal pilihan ganda terdapat 34% soal memiliki daya pembeda yang baik dengan rentang indeks (0,41 – 0,64); 40% cukup baik dengan rentang indeks (0,21 – 0,38); 13% jelek dengan rentang indeks (0,05 – 0,13); dan 13% sangat jelek dengan rentang indeks (-0,10 – 0,00). Soal yang memiliki daya pembeda yang jelek bahkan negatif mengindikasikan bahwa perlu adanya perbaikan untuk diketahui lebih lanjut penyebabnya, sedangkan soal dengan daya pembeda yang cukup baik dan baik berarti sudah mampu membedakan antara siswa yang sudah dan belum memahami materi dengan baik.

4. Tingkat Kesukaran

Hasil analisis butir soal objektif PSAS Gasal Ekonomi menggunakan *software* Anates V.4.09 pada aspek tingkat kesukaran dapat dibedakan menjadi tiga kategori yaitu mudah, sedang, dan sukar. Hasil penelitian pada tingkat kesukaran butir soal objektif PSAS Gasal Ekonomi menggunakan *software* Anates V.4.09, tertera pada tabel berikut.

Tabel 3. Distribusi Tingkat Kesukaran Soal PSAS

Kategori Soal	Nomor Soal	Persentase (Jumlah Soal)
Mudah ($P > 0,70$)	15,20,23,28,30	17% (5 butir)
Sedang ($0,30 \leq P \leq 0,70$)	2,3,4,5,8,9,10, 11,13,16,17,21, 22,24,25,27,29	56% (17 butir)
Sukar ($P < 0,30$)	1,6,7,12,14,18,19,26	27% (8 butir)

Sumber: Data Diolah Peneliti, 2025

Hasil analisis Tingkat kesukaran butir soal objektif PSAS Gasal Ekonomi kelas XI SMAN 1 Bumiayu berdasarkan pada kriteria yang ada menunjukkan bahwa dari 30 butir soal pilihan ganda terdapat 17% soal berada pada kategori mudah dengan rentang indeks (0,71 – 0,78); 56% berada pada kategori sedang dengan rentang indeks (0,30 – 0,65); dan 27% berada pada kategori sukar dengan rentang indeks (0,07 – 0,25). Soal dikatakan baik ketika soal tersebut tidak terlalu sukar dan tidak terlalu mudah. Hasil penelitian menunjukkan bahwa mayoritas soal soal

berada pada kategori sedang sebanyak 17 butir soal dengan proporsi ideal yang masih perlu dianalisis lebih lanjut.

5. Efektivitas Pengecoh

Hasil analisis butir soal objektif PSAS Ekonomi menggunakan *software* Anates V.4.09 pada aspek efektifitas pengecoh dalam penelitian ini dibedakan menjadi 5 kategori yaitu sangat tinggi, tinggi, cukup, rendah, dan sangat rendah. Hasil penelitian terkait efektivitas pengecoh soal PSAS Gasal Ekonomi menggunakan *software* Anates V.4.09, tertera pada tabel berikut.

Tabel 4. Distribusi Efektivitas Pengecoh Soal

Kategori Soal	Nomor Soal	Persentase (Jumlah Soal)
Sangat Baik (4 pengecoh berfungsi)	3,6,13,14, 15, 24,28	23% (7 butir)
Baik (3 pengecoh berfungsi)	2,5,8,10,11,17,21,2, 2, 23,27,29	37% (11 butir)
Cukup (2 pengecoh berfungsi)	1,4,7,9,12,16,19,25, 26	30% (9 butir)
Kurang Baik (1 pengecoh berfungsi)	18,30	7% (2 butir)
Tidak Baik (0 pengecoh berfungsi)	20	3% (1 butir)

Sumber: Data Diolah Peneliti, 2025

Hasil analisis efektivitas pengecoh butir soal objektif PSAS Gasal Ekonomi kelas XI SMAN 1 Bumiayu berdasarkan pada kriteria yang ada menunjukkan bahwa dari 30 butir soal pilihan ganda terdapat 23% soal memiliki efektivitas pengecoh sangat baik karena semua pengecoh berfungsi; 37% soal dengan efektivitas pengecoh yang baik karena ada 3 pengecoh berfungsi; 30% soal dengan efektivitas pengecoh yang cukup baik karena ada 2 pengecoh berfungsi; 7% soal dengan efektivitas pengecoh yang kurang baik karena hanya ada 1 pengecoh berfungsi; dan 3% soal dengan efektivitas pengecoh yang tidak baik karena semua pengecoh tidak berfungsi. Dengan kata lain, 90% soal sudah memiliki efektifitas pengecoh yang cukup baik, baik, dan sangat baik.

B. Pembahasan

1. Analisis Validitas

Suatu tes disebut valid jika tes tersebut dengan tepat dapat mengukur apa yang hendak diukur. Validitas digunakan untuk mengukur tingkat keakuratan suatu tes yang berfungsi sebagai alat ukur (Fiska et al., 2021).

Hasil analisis menunjukkan bahwa validitas butir soal PSAS Gasal Ekonomi kelas XI SMAN 1 Bumiayu tahun ajaran 2024/2025 tergolong rendah, dengan hanya 23% soal memenuhi kriteria validitas. Kondisi ini menunjukkan adanya kelemahan dalam keterkaitan soal dengan indikator kompetensi yang seharusnya diukur.

Sejalan dengan penelitian (Ikawati et al., 2024) bahwa butir soal dengan validitas soal yang rendah dalam hal ini soal yang tidak valid perlu dilakukan perbaikan. Rendahnya validitas butir soal menunjukkan bahwa sebagian besar soal tidak secara akurat mengukur kompetensi yang ditargetkan. Jika dibiarkan, kondisi ini berpotensi menyesatkan dalam pengambilan keputusan pembelajaran seperti remedial, penguatan konsep, hingga penilaian akhir semester.

Validitas soal yang rendah dimungkinkan adanya ketidakselarasan dengan indikator kompetensi yang diukur, juga disebabkan karena rumusan kompleksitas materi, atau karena kurang sesuai dengan kemampuan siswa. Sementara butir soal yang valid menggambarkan bahwa soal sudah tepat dan akurat mengukur kemampuan siswa. Hal ini sesuai dengan pendapat (...) yang mengemukakan bahwa soal buatan guru pada umumnya memang belum melalui proses validitas, berbeda dengan soal terstandar yang sudah diujikan sebelumnya didukung dengan permasalahan yang dialami oleh guru yang memiliki keterbatasan waktu karena beban kerja yang cukup banyak.

2. Analisis Reliabilitas

Reliabilitas mengukur sejauh mana suatu item tes dapat menghasilkan hasil yang sama dalam berbagai percobaan berulang. Suatu tes yang mempunyai taraf kepercayaan tinggi jika mampu memberikan hasil yang konsisten/tetap/ajeg (Alista & Syahzanani, 2023).

Berdasarkan informasi hasil penelitian, butir soal objektif PSAS Gasal Ekonomi kelas XI SMAN 1 Bumiayu tahun ajaran 2024/2025 menunjukkan bahwa tingkat reliabilitas tes berada pada kategori cukup reliabel dalam mengukur kemampuan siswa. (Rahmasari & Ismiyati, 2016) menjelaskan bahwa beberapa faktor yang memengaruhi nilai reliabilitas meliputi penyelenggaraan tes, jumlah peserta tes, dan kesukaran butir soal. Penyelenggaraan tes seperti halnya petunjuk pengerjaan, kurangnya kesiapan siswa dalam menjawab soal, dan tempat tes. Apabila jumlah peserta tes semakin banyak maka dapat menyebabkan banyaknya variasi jawaban. Hal tersebut akan mempengaruhi konsistensi dari butir soal. Lebih lanjut, soal sukar yang terlalu banyak juga berakibat pada rendahnya nilai reliabilitas. karena siswa akhirnya menebak jawaban karena putus asa ketika sedang mengerjakan (Anita et al., 2018).

3. Analisis Daya Pembeda

Analisis daya pembeda soal berkaitan dengan kemampuan untuk membedakan siswa dengan kemampuan rendah (asor) dan tinggi (unggul) serta memperjelas kemampuan butir soal yang diujikan berdasarkan perolehan skor setiap siswa. Pada Anates, terdapat pembagian pada kelompok unggul dan kelompok asor yang menerapkan pengambilan sampel 27% untuk masing-masing kelompok. Dari jumlah sampel penelitian sebanyak 143 siswa, maka sebanyak masing-masing kelompok unggul dan kelompok asor terpilih 39 siswa.

Berdasarkan pemaparan hasil penelitian menunjukkan bahwa butir soal sebesar 74% sudah mempunyai daya beda pada kategori cukup baik dan baik. Dengan kata lain, sebanyak 74% butir soal objektif PSAS Gasal Ekonomi kelas XI SMAN 1 Bumiayu tahun ajaran 2024/2025 dapat dikatakan mampu membedakan antara siswa yang belum memahami dan siswa yang sudah memahami materi dengan baik. Sedangkan masih terdapat 26% butir soal yang memiliki daya pembeda jelek dan sangat jelek. Temuan penelitian ini melihat bahwa daya pembeda jelek karena soal tersebut dalam kategori terlalu sukar atau mudah. Sejalan dengan penelitian (Purniasari et al., 2021) salah satu

penyebab buruknya daya pembeda suatu soal karena soal tersebut kemungkinan terlalu sukar atau mudah, sehingga antara siswa yang tidak paham dan siswa yang sudah paham tentang materi menjadi tidak dapat dibedakan.

4. Analisis Tingkat Kesukaran

Tingkat kesukaran butir soal digunakan untuk menentukan perbandingan antara soal mudah, sedang dan sukar. Soal yang dibuat sebaiknya tidak terlalu mudah ataupun terlalu sukar. Jika terlalu mudah, membuat siswa tidak termotivasi untuk berusaha lebih keras memecahkan soal. Namun, jika terlalu sukar, akan membuat siswa putus asa atau tidak semangat karena tidak terjangkau oleh siswa. Menurut Sudjana, tingkat kesukaran soal dapat menggunakan perbandingan 1:1:1, 3:4:3, atau 3:5:2 (Ambarwati & Ismiyati, 2021)

Berdasarkan pemaparan hasil penelitian, menunjukkan bahwa 5 butir soal (17%) termasuk soal yang mudah, 17 butir soal (56%) termasuk soal sedang, dan 8 butir soal (27%) termasuk soal sukar. Tingkat kesukaran soal objektif PSAS Gasal Ekonomi kelas XI SMAN 1 Bumiayu tahun ajaran 2024/2025 dapat dikatakan belum sesuai dengan proporsi ideal. Jika dibagi dengan proporsi 1:1:1, maka seharusnya dari 30 butir soal akan ada 10 soal baik kategori mudah, sedang, maupun sukar. Jika menggunakan proporsi 3:4:3, maka akan ada 9 soal mudah, 12 soal sedang, dan 9 soal sukar. Sementara itu, jika menggunakan proporsi 3:5:2, maka pembagiannya menjadi 9 soal mudah, 15 soal sedang, dan 6 soal sukar. Meskipun hasil penelitian ini belum seimbang, soal dengan kategori sedang tetap memiliki jumlah terbanyak, sehingga tidak didominasi soal yang terlalu sukar ataupun soal yang terlalu mudah. Butir soal yang terlalu sukar dapat membuat siswa frustrasi, putus asa, dan kehilangan motivasi dalam mengerjakan sedangkan soal yang terlalu mudah dapat menyebabkan siswa tidak tertantang untuk berpikir lebih dalam sehingga berdampak pada menurunnya minat pada pelajaran dan kemampuan berpikir yang kurang terasah.

5. Analisis Efektivitas Pengecoh

Pada *software* Anates V.4.09, setiap pilihan jawaban pada butir soal dapat dikatakan berfungsi apabila bertanda (++) sangat baik, (+) baik, dan (-) kurang baik, sedangkan pilihan jawaban dapat dikatakan tidak berfungsi apabila bertanda (--) buruk dan (---) sangat buruk serta terdapat blok berwarna merah yang mengindikasikan bahwa alternatif jawaban tersebut dinyatakan tidak berfungsi. Jika dianalisis lebih lanjut, setiap butir soal pilihan ganda pada penelitian ini memiliki 5 opsi jawaban (A, B, C, D, E) yang berarti akan ada 4 pengecoh soal di setiap butir soalnya. Dari keseluruhan pengecoh sebanyak 120 opsi, terdapat 81 (67,5%) pengecoh yang berfungsi (sangat baik, baik, dan kurang baik) sedangkan sisanya 39 (32,5%) pengecoh tidak berfungsi (buruk dan sangat buruk).

Pengecoh/distraktor dikatakan efektif artinya mampu mengelabui siswa yang tidak tahu kunci jawaban melalui pemberian opsi yang salah. Semakin banyak siswa memilih pengecoh berarti keberfungsiannya semakin baik terutama jika dipilih oleh siswa dari kelompok asor. Soal dengan pilihan pengecoh yang sangat baik, baik, dan cukup baik memiliki opsi jawaban yang berfungsi baik dan dipilih secara merata oleh siswa atau peserta tes. Sedangkan soal dengan pilihan pengecoh yang kurang baik dan tidak baik perlu dikaji lebih lanjut untuk diketahui penyebabnya karena tidak berfungsi dengan baik. Pilihan jawaban yang tidak seragam, susunan kata pada soal, dan kunci jawaban yang kurang tepat dapat menjadi faktor pendukung pemilihan distraktor (Karindi & Ruffi'i, 2024). Keberhasilan pengecoh juga dapat terjadi ketika siswa yang tidak terlalu memahami materi sehingga memilih pengecoh tersebut.

Distraktor yang terbukti tidak efektif menunjukkan bahwa opsi jawaban tidak menarik perhatian siswa yang tidak memahami materi. Hal ini bisa disebabkan oleh redaksi pengecoh yang sudah terlalu jelas salah, atau terlalu mirip dengan kunci. Guru perlu mengevaluasi formulasi opsi agar setara dalam daya tarik, khususnya dengan tetap memerhatikan tingkatan berpikir dan miskonsepsi umum siswa.

6. Analisis Keseluruhan Kualitas Butir Soal

Butir soal yang telah dianalisis baik dari aspek validitas, reliabilitas, daya pembeda, tingkat kesukaran, serta efektivitas pengecoh selanjutnya perlu dianalisis kualitasnya secara keseluruhan. Untuk mengetahui apakah butir soal layak digunakan kembali, diperbaiki, atau dibuang, maka dapat dikelompokkan berdasarkan kriteria tertentu. Kualitas butir soal secara keseluruhan dapat dikategorikan dalam 3 jenis yakni butir soal baik, kurang baik, dan tidak baik (Mulyani et al., 2022) dengan penjelasan sebagai berikut.

Kualitas Butir Soal	Nomor Soal	Tindak lanjut
Baik dan diterima (≥ 3 kriteria terpenuhi)	1,2,3,4,5,6,7,9, 10,13,16,17,20,21,22,23,24,27,28,29	67% (20 butir)
Kurang Baik dan perlu direvisi (2 kriteria terpenuhi)	11,12,14,18,26,30	20% (6 butir)
Tidak Baik dan perlu dibuang (1 kriteria terpenuhi)	8,15,19,25	13% (4 butir)

Sumber: Data Diolah Peneliti, 2025

Pada soal nomor 9, 10, 16, 17, 20, 28, dan 29, semua aspek terpenuhi dari validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh yang baik. Sementara pada soal nomor 18 dan 30 yang memerlukan perbaikan, disebabkan karena pengecoh yang kurang berfungsi dengan baik dan tingkat kesukaran yang masih terlalu sukar atau terlalu mudah. Pada soal nomor 8, 15, 19, dan 25 tergolong tidak baik karena soal tersebut memiliki daya pembeda yang sangat jelek dengan nilai negatif, soal tidak valid, dan soal dianggap terlalu sukar atau terlalu mudah meskipun sudah memiliki pengecoh yang efektif. Oleh karena itu, penggantian soal atau perombakan total struktur dan redaksi soal dibutuhkan sebelum dipertimbangkan untuk digunakan kembali pada soal-soal yang kualitasnya tidak baik.

IV. SIMPULAN DAN SARAN

A. Simpulan

Berdasarkan hasil dan pembahasan pada penelitian ini, dapat disimpulkan bahwa butir soal objektif Penilaian Sumatif Akhir Semester (PSAS) Gasal mata pelajaran Ekonomi kelas XI SMAN 1 Bumiayu tahun ajaran 2024/2025 memiliki reliabilitas butir soal pada kategori "cukup reliabel" dengan keefisien sebesar 0,52. Hasil penelitian 30 butir soal pilihan ganda, menunjukkan bahwa validitas soal masih rendah, hanya 23% butir soal yang valid. Daya pembeda soal menunjukkan bahwa 74% butir termasuk dalam kategori cukup baik, baik, dan sangat baik dalam membedakan kemampuan siswa yang sudah dan belum memahami materi dengan baik. Tingkat kesukaran soal belum sesuai dengan proporsi yang ideal, namun tetap didominasi oleh butir soal dengan kategori sedang sebanyak 56% sehingga dapat dikatakan bahwa sebagian besar soal seharusnya tidak membuat siswa putus asa karena terlalu sukar dan tidak membuat siswa kurang tertantang berpikir lebih dalam karena terlalu mudah. Efektivitas pengecoh dalam butir soal menunjukkan bahwa 90% soal dalam kategori sangat baik, baik, dan cukup baik dalam mengelabui siswa, khususnya kelompok asor dan setiap opsi dipilih secara merata oleh siswa.

Berdasarkan analisis secara menyeluruh, terdapat 67% butir soal yang berkualitas baik, 20% kurang baik, dan 13% tidak baik. Untuk butir soal yang sudah baik dapat disimpan dalam bank soal untuk digunakan kembali pada penilaian berikutnya. Sementara, butir soal yang kurang baik perlu ada perbaikan pada aspek yang tidak terpenuhi agar bisa diterima, dan butir soal yang tidak baik disarankan diganti atau diperlukan revisi yang sangat kompleks. Secara keseluruhan, hasil penelitian ini menunjukkan bahwa sebagian besar butir soal dapat diterima namun beberapa butir soal masih memerlukan perbaikan agar dapat menjadi alat ukur yang valid, reliabel, dan efektif dalam mengevaluasi pencapaian kompetensi siswa.

B. Saran

Soal yang kurang berkualitas hendaknya diperbaiki sedangkan untuk soal yang tidak berkualitas disarankan agar tidak digunakan lagi. Jika daya beda soal baik namun pengecoh kurang berfungsi, maka yang perlu diperbaiki adalah alternatif jawaban (pengecohnya),

bukan keseluruhan soal. Namun, jika daya beda bernilai negatif (sangat jelek), maka butir soal sebaiknya dibuang.

Penelitian ini terbatas pada kedalaman analisis karena fokus penelitian ini pada deskripsi hasil analisis dari data penelitian menggunakan pendekatan kuantitatif. Maka, disarankan bagi peneliti selanjutnya untuk mengkaji lebih mendalam terkait faktor-faktor penyebab dari setiap aspek yang dianalisis atau menambahkan pendekatan kualitatif dalam teknik analisisnya untuk mendapatkan hasil penelitian yang lebih memuaskan dan komprehensif. Penggunaan *software* lain sebagai alat analisis juga disarankan untuk mengeksplorasi dan membandingkan hasil.

DAFTAR RUJUKAN

- Adom, D., Mensah, J. A., & Dake, D. A. (2020). Test , measurement, and evaluation : Understanding and use of the concepts in education. *International Journal of Evaluation and Research in Education (IJERE)*, 9(1), 109-119. <https://doi.org/10.11591/ijere.v9i1.20457>
- Alista, Y. F., & Syahzanani, R. A. (2023). Analisis Butir Soal Ulangan Harian Fisika dengan Pendekatan Teori Tes Klasik Menggunakan Program Anates. *SIPTEK : Seminar Nasional Inovasi Dan Pengembangan Teknologi Pendidikan*, 1(1), 1-11. <https://proceeding.unesa.ac.id/index.php/siptek/article/view/173>
- Alti, R. P., Zulyusri, & Violita. (2022). Analisis Kualitas Butir Soal Ujian Tengah Semester Mata Pelajaran Biologi Kelas X di MAN 1 Solok Selatan. *Jurnal Metaedukasi: Jurnal Ilmiah Pendidikan*, 4(2), 70-75. <https://doi.org/10.37058/metaedukasi.v4i2.4089>
- Ambarwati, Y. F., & Ismiyati, I. (2021). Analisis Butir Soal Pilihan Ganda Ulangan Akhir Semester Genap Mata Pelajaran Kearsipan. *Measurement in Educational Research (Meter)*, 1(2), 64-75. <https://doi.org/10.33292/meter.v1i2.144>
- Anita, Tyowati, S., & Zulfadrial. (2018). Analisis Kualitas Butir Soal Fisika Kelas X Sekolah Menengah Atas. *Jurnal Pendidikan Bahasa Indonesia*, 16(1), 35-47. <https://doi.org/10.31571/edukasi.v16i1.780>

- Augustia, A. D., Augustia, C. N., Azzahra, D., Hakim, L., & Pratiwi, V. (2025). Analisis Validitas dan Reliabilitas Soal Pilihan Ganda dengan Menggunakan Software Anates pada Mata Pelajaran Perpajakan. *AKADEMIK: Jurnal Mahasiswa Ekonomi & Bisnis*, 5(1), 250–265.
<https://doi.org/10.37481/jmeb.v5i1.1165>
- Cahyani, D. A., Aurelia, A. H., Hakim, L., & Pratiwi, V. (2024). Analisis Butir Soal HOTS Akuntansi Perusahaan Jasa Menggunakan Software Anates pada Siswa SMK Kelas 11. *PESHUM: Jurnal Pendidikan, Sosial Dan Humaniora*, 4(1), 1334–1346.
<https://doi.org/10.56799/peshum.v4i1.7108>
- Farida, & Musyarofah, A. (2021). Validitas dan Reliabilitas dalam Analisis Butir Soal. *Al-Mu'arrif: Journal of Arabic Educatioan*, 1(1), 34–44. <https://doi.org/10.32923/al-muarrib.v1i1.2100>
- Fietri, W. A., Zulyusri, & Violita. (2021). Analisis Butir Soal Biologi Kelas XI Madrasah Aliyah Sakinah Kerinci Menggunakan Program Komputer Anates 4.0 For Windows. *Natural Science: Jurnal Penelitian Bidang IPA Dan Pendidikan IPA*, 7(1), 28–35.
<https://doi.org/10.15548/nsc.v7i1.2329>
- Fiska, J. M., Hidayati, Y., Qomaria, N., & Hadi, W. P. (2021). Analisis Butir Soal Ulangan Harian IPA Menggunakan Software Anates pada Pendekatan Teori Tes Klasik. *Jurnal Natural Science Educational Research*, 4(1).
<https://doi.org/10.21107/nser.v4i1.8133>
- Hikmawati, F. (2020). *Metodologi Penelitian* (1 cetakan 4). Depok:Rajawali Pres PT RajaGrafindo Persada.
- Ikawati, D. S., Agustin, E. W., Hakim, L., & Pratiwi, V. (2024). Analisis Butir Soal Pilihan Ganda pada Siswa SMK Menggunakan Anates. *Edukatif: Jurnal Ilmu Pendidikan*, 6(6), 6773–6781.
- Kaka, L., Bano, V. O., & Njoeroemana, Y. (2024). Efektivitas Analisis Butir Soal Pilihan Ganda Menggunakan Aplikasi Anates Di SMPN 2 Kanatang. *JIP: Jurnal Inovasi Penelitian*, 4(9), 1441–1450.
<https://doi.org/10.47492/jip.v4i9.3124>
- Karindi, D. R., & Rufi'i. (2024). Penggunaan ANATES 4.0.9 dalam Menguji Kelayakan Soal Pilihan Ganda Sumatif Matematika. *LINEAR: Journal of Mathematics Edukation*, 5(2). <https://doi.org/10.32332/2zg4f753>
- Mania, S., Fitriani, Majid, A. F., Ichiana, N. N., & Abrar, A. I. P. (2020). Analisis Butir Soal Ujian Akhir Sekolah. *Al Asma: Journal of Islamic Education*, 2(2), 274–284.
<https://doi.org/10.24252/asma.v2i2.16569>
- Mulyani, S., Krismonita, M., & Yamtinah, S. (2022). Analisis Butir Soal dan Kecukupan HOTS Soal Ujian Akhir Semester Mata Pelajaran Kimia Kelas X. *Paedagogia: Jurnal Penelitian Pendidikan*, 25(2).
<https://doi.org/10.20961/paedagogia.v25i2.60913>
- Ningsih, S. N., Istiyono, E., & Widiastuti. (2023). An Application of Classical Test Theory for Item Characteristic Analysis of Chemical Literacy Instruments. *Jurnal Pendidikan Kimia Indonesia*, 7(2).
<https://doi.org/10.23887/jpki.v7i2.68983>
- Nirwana, R., Hidayati, A. I., Ifcha, F. A., Azzahra, S. F., & Jannah, A. S. R. (2024). Penilaian dalam Kurikulum Merdeka: Mendukung Pembelajaran Adaptif dan Berpusat pada Siswa Madrasah Ibtidaiyah. *Jurnal Madrasah Ibtidaiyah (JMI)*, 02(02), 213–224.
<https://journal.unusida.ac.id/index.php/jmi/article/view/1320>
- Nurhalimah, S., Selfa, S. K., Amanda, U., & Ilmi, D. (2024). Penguatan Pendidikan Peserta Didik Melalui Keterampilan 4C dengan Mengimplementasikan Pembelajaran Abad 21 dalam Pembelajaran Pendidikan Agama Islam. *Al-Tarbiyah: Jurnal Ilmu Pendidikan Islam*, 2(1). <https://doi.org/10.59059/al-tarbiyah.v2i1.799>
- Purniasari, L., Masykuri, M., & Ariani, S. R. D. (2021). Analisis Butir Soal Ujian Sekolah Mata Pelajaran Kimia SMA N 1 Kutowinangun Tahun Pelajaran 2019 / 2020 Menggunakan Model Itean dan Rasch. *Jurnal Pendidikan Kimia*, 10(2).
<https://doi.org/10.20961/jpkim.v10i2.48244>
- Rahmasari, D., & Ismiyati. (2016). Analisis Butir Soal Mata Pelajaran Pengantar Administrasi Perkantoran. *Economic Education Analysis Journal (EEAJ)*, 5(1),

317–330.

<https://journal.unnes.ac.id/sju/eeaj/article/view/10007>

- Riani, D., Almujab, S., Dina, A., Fitriani, & Budiarto, R. (2020). Analisis Butir Soal dan Kemampuan Siswa dalam Menjawab Soal Ujian Nasional pada Mata Pelajaran Ekonomi. *Oikos: Jurnal Kajian Pendidikan Ekonomi Dan Ilmu Ekonomi*, *IV*(1), 70–79. <https://doi.org/10.23969/oikos.v4i1.2425>
- Saely, E., & Shaleh. (2023). Analisa Butir Soal Penilaian Akhir Semester (PAS) Pada Mata Pelajaran PPKN Menggunakan Aplikasi Winstep dan Item And Analysis (ITEMAN) 4.3. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, *08*(03), 415–430. <http://journal.unpas.ac.id/index.php/pendas/article/view/10613>
- Sari, P. I., & Yudha, R. I. (2020). Pemanfaatan Penerapan Media Berbasis Software Anates pada Mata Kuliah Evaluasi Pembelajaran di Universitas Batanghari Jambi. *Jurnal Ilmiah Universitas Batanghari Jambi (JIUBJ)*, *20*(1), 81–85. <https://doi.org/10.33087/jiubj.v20i1.732>
- Umami, R., & Rusdi, M. (2021). Pengembangan instrumen tes untuk mengukur Higher Order Thinking Skills (HOTS) berorientasi Programme for International Student Assessment (PISA) pada peserta didik. *JP3M: Jurnal Penelitian Pendidikan Dan Pengajaran Matematika*, *7*(1), 57–68. <https://doi.org/10.37058/jp3m.v7i1.2069>
- Walfath, I., & Faurenza, R. (2024). Analisis dan Evaluasi Instrumen dalam Pembelajaran Matematika pada Materi Pecahan di Sekolah Dasar. *JEDMA: Jurnal Edukasi Matematika*, *4*(2). <https://doi.org/10.51836/jedma.v4i2.627>
- Wulandari, O., Muhtarom, M., & Sumarno, S. (2025). Analisis Butir Soal Pengetahuan Matematika Kelas V Sekolah Dasar Menggunakan Model Rasch. *Jurnal Pengembangan Dan Penelitian Pendidikan*, *07*(1), 293–302. <https://journalpedia.com/1/index.php/jpp/article/view/4435>
- Zalukhu, K. O., Khairuddin, Syamsuar, & Sari, D. N. (2023). Evaluasi Instrumen Pembelajaran Pendidikan Jasmani Olahraga dan Kesehatan Kelas VII Sekolah Menengah Pertama Negeri 30 Padang. *JPDO: Jurnal Pendidikan Dan Olahraga*, *6*(6), 89–96. <http://jpdo.ppj.unp.ac.id/index.php/jpdo/article/view/1407>