



Analisis Kualitas Butir Soal Pilihan Ganda Elemen Dokumen Berbasis Digital Menggunakan ANATES

Betty Sofi Hariati ✉, Universitas Negeri Surabaya

Luqman Hakim, Universitas Negeri Surabaya

Vivi Pratiwi, Universitas Negeri Surabaya

Ainuz Zakia, Universitas Negeri Surabaya

Revalina Cintia Ayu, Universitas Negeri Surabaya

Andrea Dzakira Maheswari, Universitas Negeri Surabaya

Shafira Eka Salsabila, Universitas Negeri Surabaya

Qiptiyah Qurrotul Aini, Universitas Negeri Surabaya

Tri Dwi Rahmawati, Universitas Negeri Surabaya

✉ bettyhariati3@gmail.com

Abstract: This study was conducted to evaluate the quality of multiple-choice questions related to Digital Document Elements using ANATES version 4. The main problem the uncertainty about how good the quality of each question was based on the level of difficulty, discriminating ability, effectiveness of incorrect choices, and validity seen from the results. This study used quantitative methods and descriptive analysis on 63 questions that had been answered by 20 participants. Based on the results of the study, the tool used showed a very high level of reliability and could be used for assessment. However, most of the questions were classified as very easy and easy, so they were not fully effective in measuring students' abilities. In the discrimination power analysis, only a few questions showed good discrimination ability. In addition, analysis of the incorrect choices revealed several choices that did not work well. These results are an important foundation for developing better, more effective evaluation tools that support the improvement of learning quality.

Keywords: Item analysis, ANATES, Difficulty level, Discrimination index

Abstrak: Penelitian ini dilakukan untuk mengevaluasi kualitas soal pilihan ganda yang berkaitan dengan Elemen Dokumen Berbasis Digital dengan memanfaatkan ANATES versi 4. Masalah utama ketidakpastian tentang seberapa baik kualitas setiap soal berdasarkan tingkat kesulitan, kemampuan untuk membedakan, efektivitas pilihan yang salah, dan validitas yang terlihat dari hasil. Penelitian ini menggunakan metode kuantitatif dan analisis deskriptif pada 63 soal yang sudah dijawab oleh 20 peserta. Berdasarkan hasil penelitian, alat yang digunakan menunjukkan tingkat keandalan yang sangat tinggi dan bisa dipakai untuk melakukan penilaian. Namun, sebagian besar soal nya tergolong sangat mudah dan mudah, sehingga belum sepenuhnya efektif dalam mengukur kemampuan siswa.. Dalam analisis daya pembeda, hanya beberapa soal yang menunjukkan kemampuan membedakan yang baik. Selain itu, analisis terhadap pilihan yang salah mengungkapkan beberapa pilihan yang tidak bekerja dengan baik. Hasil ini menjadi fondasi penting untuk mengembangkan alat evaluasi lebih baik, efektif, dan mendukung peningkatan kualitas pembelajaran.

Kata kunci: Analisis Butir Soal, ANATES, Tingkat Kesukaran, Daya Pembeda

Received 20 November 2025; **Accepted** 30 Januari 2026; **Published** 10 Februari 2026

Citation: Hariati, B.S., Hakim, L., Pratiwi, V., Zakia, A., Ayu, R.C., Maheswari, A.D., Salsabila, S.E., Aini, Q.Q., Rahmawati, T.D. (2026). Analisis Kualitas Butir Soal Pilihan Ganda Elemen Dokumen Berbasis Digital Menggunakan ANATES. *Jurnal Jendela Pendidikan*, 6 (01), 46-58.



Copyright ©2026 Jurnal Jendela Pendidikan

Published by CV. Jendela Edukasi Indonesia. This work is licensed under the Creative Commons Attribution-NonCommercial-Share Alike 4.0 International License.

PENDAHULUAN

Pada dekade terakhir, kemajuan teknologi digital telah mengubah banyak aspek kehidupan, termasuk pendidikan. Transformasi digital tidak lagi hanya menjadi tambahan untuk pembelajaran, tetapi telah menjadi bagian penting dari manajemen informasi, distribusi materi, dan evaluasi pembelajaran. Pada era pembelajaran modern, literasi digital, kemampuan berpikir kritis, dan kemampuan pemecahan masalah berbasis informasi adalah keterampilan yang sangat penting bagi siswa untuk dikuasai. Alat evaluasi yang bukan hanya dapat dipercaya dan dapat diandalkan, akan tetapi juga dapat memberi gambaran yang akurat dan menyeluruh tentang kemampuan siswa diperlukan untuk mendukung pencapaian kompetensi tersebut. Karena dianggap efektif untuk mengukur berbagai tingkat kemampuan kognitif dalam cakupan materi yang luas, soal objektif menjadi salah satu alat evaluasi yang sering dipilih di berbagai jenjang pendidikan. Namun demikian, kualitas instrumen pilihan ganda masih menjadi masalah besar, terutama dalam hal kompetensi digital seperti materi Elemen Dokumen Berbasis Digital. Banyak alat soal yang digunakan di sekolah masih belum melalui proses analisis detail menyeluruh, sehingga kualitasnya tidak dapat dijamin sepenuhnya (Astuti & Mulyani, 2023; Hapsari & Widodo, 2023).

Problem utama dengan soal pilihan ganda adalah ketidaktahuan kualitas masing-masing soal. Ini terutama berlaku untuk tingkat kesukaran, daya pembeda, efektivitas pengecoh, dan validitas soal. Keempat elemen itu sangat penting untuk menentukan apakah alat evaluasi dapat memberikan hasil penilaian yang representatif dan dapat dipercaya (Nasyrah, Bustanur, & Zulhaini, 2023). Studi sebelumnya menemukan bahwa instrumen evaluasi sering mengandung butir yang terlalu mudah, sehingga tidak efektif dalam mengidentifikasi antara peserta didik berkemampuan tinggi atau rendah, dan memiliki pengecoh yang tidak berfungsi (Mustaqim & Sulisti, 2024). Dalam situasi seperti ini, hasil penilaian menjadi tidak akurat dan tidak mencerminkan kemampuan sebenarnya siswa. Kualitas soal menjadi semakin penting dalam pembelajaran digital karena materi cenderung memerlukan pemahaman konseptual dan keterampilan teknis untuk menggunakan teknologi. Proses penilaian tidak dapat menunjukkan tingkat penguasaan siswa terhadap materi jika alat evaluasi tidak dianalisis secara menyeluruh (Masruri et al., 2024).

Solusi yang efektif untuk memecahkan masalah ini adalah dengan mengaplikasikan perangkat lunak analisis butir soal seperti ANATES versi 4. ANATES dibuat untuk membantu guru dan peneliti secara cepat dan sistematis menganalisis kualitas butir soal menggunakan teori tes klasik (Astuti & Mulyani, 2023). Aplikasi ini dapat menampilkan informasi detail mengenai validitas butir, reliabilitas instrumen, tingkat kesulitan, daya pembeda, dan efektivitas pengecoh. Aplikasi anates lebih efektif daripada analisis manual, mengurangi kesalahan perhitungan, dan lebih objektif daripada analisis manual. Selain itu, ANATES menawarkan visualisasi numerik, tabel, dan grafik yang membantu pendidik untuk memahami secara menyeluruh karakteristik soal. Penelitian terbaru menunjukkan bahwa menggunakan ANATES meningkatkan kualitas instrumen evaluasi, terutama dalam pembelajaran berbasis digital yang membutuhkan ketelitian tinggi untuk analisis (Masruri et al., 2024).

Teori evaluasi pendidikan, teori analisis butir klasik, konsep reliabilitas instrumen, dan teori efektivitas pengecoh adalah beberapa teori yang mendasari penelitian ini. Sejauh mana instrumen tes mampu mengukur kemampuan siswa secara konsisten dan tepat dalam teori evaluasi pendidikan menentukan kualitasnya (Andryan Fitryansyah, Hilmiyati, & Habudin, 2024). Karena karakteristik setiap komponen sangat memengaruhi kualitas keseluruhan, teori analisis butir menyatakan bahwa setiap komponen harus dievaluasi secara empiris. Dua indikator penting yang menunjukkan kemampuan diagnostik suatu instrumen adalah tingkat kesukaran dan daya pembeda. Sementara daya pembeda menunjukkan kompetensi soal dalam menyeleksi siswa sesuai kemampuan mereka, soal

yang terlalu sederhana atau terlalu kompleks mungkin tidak memberikan informasi yang cukup untuk menggambarkan kemampuan siswa (Gunawan & Asria, 2023; Andryan Fitryansyah et al., 2024). Sebaliknya, penting untuk tidak mengabaikan kualitas pengecoh karena pengecoh yang tidak berfungsi dapat menurunkan kualitas informasi dan memungkinkan siswa menebak jawaban dengan lebih mudah (Handayani, 2022)). Selain itu, literatur tentang asesmen digital menekankan bahwa penggunaan perangkat analisis berbasis komputer dapat meningkatkan efisiensi evaluasi, mengurangi subjektivitas, dan membantu guru dalam membuat keputusan berbasis data (Kumalasari, Karaman, Mustikasari, & Kurniawan, 2022).

Studi ini dilakukan dengan melihat 63 soal pilihan ganda yang digunakan dalam pembelajaran Elemen Dokumen Berbasis Digital. Penelitian ini berfokus pada pengukuran kualitas butir soal melalui ANATES versi 4. Tujuan penelitian ini untuk menganalisis tingkat kesukaran pada soal, menilai daya pembeda soal dalam mengidentifikasi peserta didik berkemampuan tinggi atau rendah, menelaah efektivitas pengecoh untuk mengetahui apakah opsi jawaban dapat bekerja semestinya, serta mengukur validitas butir dan reliabilitas instrumen secara keseluruhan berdasarkan hasil dari ANATES. Analisis tersebut diperlukan untuk membuktikan bahwa instrumen yang digunakan bisa menggambarkan kemampuan siswa secara tepat.

Kajian teoritik dan penelitian terdahulu menunjukkan bahwa instrumen evaluasi yang dianalisis menggunakan ANATES mampu menghasilkan data secara tepat dan membantu menunjang mutu pembelajaran. Penelitian Astuti dan Mulyani (2023) menemukan bahwa analisis menggunakan ANATES memberikan gambaran rinci mengenai kelemahan dan membantu melakukan revisi butir soal secara sistematis. Temuan tersebut sejalan dengan penelitian Hartono, Tenriawaru, dan Ningsih (2023) yang menunjukkan bahwa penerapan analisis kuantitatif menggunakan ANATES pada evaluasi meningkatkan mutu pembelajaran karena data yang diperoleh lebih akurat dan dapat digunakan untuk pengambilan keputusan yang lebih tepat. Dengan mengintegrasikan wawasan teori dan temuan empiris tersebut, penelitian ini dimaksudkan mampu mendukung pengembangan evaluasi pada pendidikan berbasis digital, khususnya pada Elemen Dokumen Berbasis Digital.

METODE

Penelitian ini digunakan untuk pendekatan kuantitatif deskriptif dengan fungsi mendeskripsikan dan menganalisis kualitas soal HOTS bertema “Dokumen Berbasis Digital”. Penelitian ini kami lakukan sebagai bagian dari kegiatan mata kuliah *Evaluasi Belajar dan Pembelajaran* yang berfokus pada penilaian kualitas butir soal yg kita gunakan dari aspek tingkat kesulitan, daya pembeda, reliabilitas, dan efektivitasnya dengan mengukur potensi berpikir kritis peserta didik. Subjek penelitian terdiri dari 20 siswa/siswi SMK jurusan Akuntansi yang menjadi responden dalam uji coba soal, sedangkan objek penelitian berupa 63 butir soal HOTS yang telah disusun oleh mahasiswa Prodi Pendidikan Akuntansi Universitas Negeri Surabaya. Data yang kami gunakan dalam penelitian ini mencakup hasil jawaban siswa terhadap soal HOTS serta kunci jawaban yang telah divalidasi oleh penyusun.

Teknik pengumpulan data dilakukan secara daring melalui Google Form, yang memudahkan proses distribusi dan pengumpulan hasil jawaban peserta. Metode ini dipilih untuk menyesuaikan dengan perkembangan teknologi pendidikan dan pembelajaran abad ke-21. Analisis data kami lakukan secara kuantitatif dengan menggunakan pendekatan teori tes klasik (Classical Test Theory). Analisis berdasarkan penghitungan validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh (distractor) pada setiap butir soal. Pengolahan data dilakukan dengan bantuan program Anates guna memperoleh hasil analisis yang akurat dan sistematis.

Reliabilitas dan Daya Pembeda

Reliabilitas merupakan cara menilai untuk kita bisa mengetahui suatu instrumen yang mampu kita terima dalam belum hasil konsisten ketika diujikan dalam kondisi yang relatif serupa. Instrumen yang mempunyai reliabilitas tinggi menunjukkan bahwa alat ukur tersebut dapat dipercaya untuk menilai potensi peserta secara konsisten dari waktu ke waktu. Dengan maksud lain, hasil pengukuran tidak mudah berubah meskipun dilakukan pada kesempatan yang berbeda.

TABEL 1. Kriteria Koefisien Reliabilitas

No	Koefisien Reliabilitas	Kriteria
1	0 - 54	Sangat Rendah
2	55 - 64	Rendah
3	65 - 79	Sedang
4	80 - 89	Tinggi
5	90 - 100	Sangat Tinggi

Sumber : (Taherdoost, 2024)

Sedangkan, daya pembeda digunakan untuk melihat seberapa baik suatu butir soal mampu menunjukkan perbedaan peserta didik yang memiliki kemampuan tinggi dan rendah. Semakin besar nilai daya beda, berarti suatu butir soal efektif dalam mengungkapkan variasi tingkat pemahaman materi dari peserta didik. Maka, butir soal yang memiliki daya beda tinggi dianggap efektif dan pantas digunakan dalam evaluasi pembelajaran.

TABEL 2. Klasifikasi Daya Beda

No	Klasifikasi	Keterangan
1	0,00 - 0,19	Tidak Baik (soal direvisi)
2	0,20 - 0,39	Cukup Baik (soal digunakan & direvisi)
3	0,40 - 0,69	Baik (soal di gunakan
4	0,70 - 1,00	Sangat Baik (soal digunakan)
5	Negatif	Soal Dibuang

Sumber : (Anastasi & Urbina, 2023)

Tingkat Kesukaran

Tingkat kesukaran memiliki fungsi untuk mengetahui potensi siswa saat mengerjakan suatu butir soal mudah atau sulit. Suatu butir soal yang baik umumnya memiliki tingkat kesukaran yang seimbang, yaitu tidak terlalu mudah dan tidak terlalu sukar, sehingga bisa membedakan kemampuan peserta dengan lebih objektif.

TABEL 3. Klasifikasi Tingkat Kesukaran

No	Indeks Tingkat Kesukaran	Penafsiran Butri Soal
1	$0,00 \leq P \leq 0,30$	Sukar
2	$0,30 \leq P \leq 0,70$	Sedang
3	$0,70 \leq P \leq 1,00$	Mudah

Sumber : (Puspendik, 2023)

Validitas dan Efektivitas Pengecoh

Validitas berfungsi untuk memastikan suatu instrumen penilaian benar-benar mengukur kemampuan atau aspek yang akan diukur. Suatu tes dikatakan memiliki validitas yang baik

jika hasil pengukurannya sesuai dengan tujuan penyusunan tes tersebut (Puspasari & Puspita, 2022). Semakin tinggi nilai validitas butir soal, semakin tepat pula soal tersebut dalam merepresentasikan kompetensi peserta didik. Menurut Windyastuti (2023), butir soal dengan nilai korelasi tinggi tergolong soal yang berkualitas dan layak digunakan dalam kegiatan penilaian.

TABEL 4. *Kriteria Koefisien Validitas*

No	Koefisien Butir Soal	Interpretasi
1	0,80 - 1.00	Sangat Tinggi
2	0,60 - 0,79	Tinggi
3	0,40 - 0,59	Cukup
4	0,20 - 0,39	Rendah
5	0,00 - 0,19	Sangat Rendah

Sumber : (Windyastuti, 2023)

Sedangkan, efektivitas pengecoh digunakan untuk menilai seberapa baik pilihan jawaban yang salah (distraktor) dalam mengelabui peserta yang belum memahami materi. Distraktor yang baik akan menarik perhatian siswa yang belum menguasai konsep, sehingga membantu melihat perbedaan kemampuan peserta tes. Setiap opsi jawaban dinilai berdasarkan kriteria tertentu, yaitu (++) sangat baik, (+) baik, (-) kurang baik, (--) buruk, dan (---) sangat buruk. Pilihan jawaban yang tidak berfungsi ditandai dengan blok khusus yang menunjukkan bahwa pengecoh tersebut tidak efektif.

HASIL PENELITIAN

Reliabilitas menyiratkan konsisten hasil pengukuran. Alat ukur dikatakan reliabel apabila memberikan hasil yang sama dalam kondisi serupa. Dalam literatur evaluasi pendidikan reliabilitas diartikan sebagai tingkat keterpercayaan hasil pengukuran, yaitu sejauh mana dapat diandalkan apabila pengujian ulang pada siswa yang sama dalam waktu berbeda masih menunjukkan skor yang konsisten. Selain itu, reliabilitas juga berkaitan dengan minimnya kesalahan pengukuran sehingga skor yang dihasilkan lebih stabil dan dapat dipercaya (Anshari et al., 2024).

Berdasarkan hasil pengolahan data yang dilakukan melalui program Anates, diperoleh hasil sebagai berikut.

RELIABILITAS TES					
=====					
Rata2= 57,30					
Simpang Baku= 6,33					
KorelasiXY= 0,94					
Reliabilitas Tes= 0,97					
No.Urut	No. Subyek	Kode/Nama Subyek	Skor Ganjil	Skor Genap	Skor Total
1	1	Putri Arvy Oc...	23	26	49
2	2	Irma Zuliah F...	25	26	51
3	3	Aulia Sari Ri...	24	26	50
4	4	Devita Maharani	24	23	47
5	5	Marimbie Kart...	30	30	60
6	6	Angelo Kusuma...	32	31	63
7	7	Werdeh	20	21	41
8	8	Vonny Kartika...	32	31	63
9	9	Aisyah Dita R...	32	30	62
10	10	Rayhanah Zahw...	31	30	61
11	11	Nabilatus Sho...	30	29	59
12	12	Nia Rahmawati	32	31	63
13	13	Andia Nafila	31	31	62
14	14	Dina 'Aufa	31	30	61
15	15	Florensia Ang...	29	29	58
16	16	Irnanda Purba...	31	29	60
17	17	Abiyan Hasanain	32	31	63
18	18	Revita Anjani...	28	29	57
19	19	Kaila Renata ...	30	30	60
20	20	Raeya Adeline	29	27	56

Sumber : Hasil Olah Data, 2025

GAMBAR 1. Hasil Analisis Reliabilitas

Daya pembeda merupakan kemampuan butir soal dalam membedakan peserta didik yang memiliki tingkat kemampuan tinggi dari mereka yang kemampuannya lebih rendah. Daya pembeda yang baik akan memperlihatkan bahwa peserta didik dengan kompetensi tinggi cenderung menjawab dengan benar, sedangkan peserta didik dengan kompetensi rendah sering memberikan jawaban yang salah. Dapat disimpulkan bahwa daya pembeda berperan penting dalam mengukur penguasaan materi setiap peserta didik (Fitryansyah et al., 2024; Nurhayati et al., 2024).

Berdasarkan hasil pengolahan data yang dilakukan melalui ANATES, diperoleh hasil sebagai berikut.

DAYA PEMBEDA									

Jumlah Subyek= 20									
Kip Atas/Bawah(n)= 5									
Butir Soal= 63									
No Butir Baru	No Butir Asli	Kel. Atas	Kel. Bawah	Beda	Indeks DP (%)	27	28	29	30
1	1	5	0	5	100,00	27	28	29	30
2	2	5	0	5	100,00	28	29	30	31
3	3	5	0	5	100,00	29	30	31	32
4	4	5	0	5	100,00	30	31	32	33
5	5	5	1	4	80,00	31	32	33	34
6	6	5	1	4	80,00	32	33	34	35
7	7	5	1	4	80,00	33	34	35	36
8	8	5	1	4	80,00	34	35	36	37
9	9	5	1	4	80,00	35	36	37	38
10	10	5	5	0	0,00	36	37	38	39
11	11	5	5	0	0,00	37	38	39	40
12	12	5	1	4	80,00	38	39	40	41
13	13	5	1	4	80,00	39	40	41	42
14	14	5	4	1	20,00	40	41	42	43
15	15	5	3	2	40,00	41	42	43	44
16	16	5	4	1	20,00	42	43	44	45
17	17	5	0	5	100,00	43	44	45	46
18	18	5	4	1	20,00	44	45	46	47
19	19	5	5	0	0,00	45	46	47	48
20	20	5	5	0	0,00	46	47	48	49
21	21	5	5	0	0,00	47	48	49	50
22	22	5	5	0	0,00	48	49	50	51
23	23	5	5	0	0,00	49	50	51	52
24	24	5	4	1	20,00	50	51	52	53
25	25	5	4	1	20,00	51	52	53	54
26	26	5	4	1	20,00	52	53	54	55
						53	54	55	56
						61	62	63	64
						62	63	64	65
						63	64	65	66

Sumber : Hasil Olah Data, 2025

GAMBAR 2. Analisis Daya Pembeda

Tingkat kesukaran suatu soal menggambarkan probabilitas peserta didik untuk menjawab benar pada level kemampuan tertentu dan dinyatakan dalam bentuk indeks. Indeks kesukaran diperoleh dari perbandingan jumlah peserta didik yang menjawab benar suatu butir soal dari jumlah keseluruhan peserta didik. dengan demikian, semakin banyak peserta didik yang menjawab benar, semakin tinggi nilai indeksnya yang menunjukkan bahwa butir soal tersebut kategori mudah (JRPP, 2024).

Berdasarkan informasi yang telah dikumpulkan serta analisis yang dilakukan menggunakan ANATES, hasilnya sebagai berikut:

TINGKAT KESUKARAN									

Jumlah Subyek= 20									
Butir Soal= 63									
No Butir Baru	No Butir Asli	Jml Betul	Tkt. Kesukaran(%)	Tafsiran	28	29	30	100,00	Sangat Mudah
1	1	15	75,00	Mudah	29	30	20	100,00	Sangat Mudah
2	2	15	75,00	Mudah	30	31	20	100,00	Sangat Mudah
3	3	15	75,00	Mudah	31	32	19	100,00	Sangat Mudah
4	4	15	75,00	Mudah	32	33	19	95,00	Sangat Mudah
5	5	15	75,00	Mudah	33	34	19	95,00	Sangat Mudah
6	6	15	75,00	Mudah	34	35	19	95,00	Sangat Mudah
7	7	16	80,00	Mudah	35	36	19	95,00	Sangat Mudah
8	8	16	80,00	Mudah	36	37	20	100,00	Sangat Mudah
9	9	15	75,00	Mudah	37	38	19	95,00	Sangat Mudah
10	10	20	100,00	Sangat Mudah	38	39	19	95,00	Sangat Mudah
11	11	20	100,00	Sangat Mudah	39	40	20	100,00	Sangat Mudah
12	12	15	75,00	Mudah	40	41	18	90,00	Sangat Mudah
13	13	14	70,00	Sedang	41	42	18	90,00	Sangat Mudah
14	14	19	95,00	Sangat Mudah	42	43	20	100,00	Sangat Mudah
15	15	18	90,00	Sangat Mudah	43	44	19	95,00	Sangat Mudah
16	16	19	95,00	Sangat Mudah	44	45	20	100,00	Sangat Mudah
17	17	11	55,00	Sedang	45	46	19	95,00	Sangat Mudah
18	18	18	90,00	Sangat Mudah	46	47	19	95,00	Sangat Mudah
19	19	20	100,00	Sangat Mudah	47	48	20	100,00	Sangat Mudah
20	20	20	100,00	Sangat Mudah	48	49	17	85,00	Mudah
21	21	20	100,00	Sangat Mudah	49	50	19	95,00	Sangat Mudah
22	22	19	95,00	Sangat Mudah	50	51	20	100,00	Sangat Mudah
23	23	18	90,00	Sangat Mudah	51	52	19	95,00	Sangat Mudah
24	24	16	80,00	Mudah	52	53	20	100,00	Sangat Mudah
25	25	17	85,00	Mudah	53	54	20	100,00	Sangat Mudah
26	26	17	85,00	Mudah	54	55	18	90,00	Sangat Mudah
27	27	17	85,00	Mudah	55	56	19	95,00	Sangat Mudah
					56	57	18	90,00	Sangat Mudah
					57	58	19	95,00	Sangat Mudah
					58	59	20	100,00	Sangat Mudah
					59	60	19	95,00	Sangat Mudah
					60	61	19	95,00	Sangat Mudah
					61	62	19	95,00	Sangat Mudah
					62	63	19	95,00	Sangat Mudah
					63	64	19	95,00	Sangat Mudah

Sumber : Hasil Olah Data, 2025

GAMBAR 3. Analisis Tingkat Kesukaran

ANATES melakukan evaluasi terhadap kevalidan butir soal berupa menganalisis keterkaitan antar nilai masing-masing pertanyaan dengan nilai total. Hasil dari evaluasi validitas dalam aplikasi ini dibedakan menjadi dua kategori, yakni hasil yang signifikan menunjukkan bahwa soal tersebut memenuhi kriteria, sedangkan hasil yang tidak signifikan yang menunjukkan bahwa butir tersebut tidak memenuhi kriteria. ANATES menampilkan output dari perhitungan hubungan antara skor tiap pertanyaan dengan skor keseluruhan.

Berdasarkan informasi yang berhasil dikumpulkan dan analisis yang dilakukan menggunakan ANATES, hasilnya sebagai berikut.

KORELASI SKOR BUTIR DG SKOR TOTAL							

Jumlah Subyek= 28				28	28	NAH	NAH
Butir Soal= 63				29	29	NAH	NAH
				30	30	NAH	NAH
				31	31	NAH	NAH
				32	32	0,283	Sangat Signifikan
				33	33	0,606	Sangat Signifikan
No Butir Baru	No Butir Asli	Korelasi	Signifikansi	34	34	-0,063	-
1	1	0,907	Sangat Signifikan	35	35	0,696	Sangat Signifikan
2	2	0,907	Sangat Signifikan	36	36	0,606	Sangat Signifikan
3	3	0,907	Sangat Signifikan	37	37	NAH	NAH
4	4	0,907	Sangat Signifikan	38	38	0,606	Sangat Signifikan
5	5	0,589	Sangat Signifikan	39	39	0,606	Sangat Signifikan
6	6	0,589	Sangat Signifikan	40	40	NAH	NAH
7	7	0,652	Sangat Signifikan	41	41	0,151	-
8	8	0,652	Sangat Signifikan	42	42	-0,119	-
9	9	0,570	Sangat Signifikan	43	43	NAH	NAH
10	10	NAH	NAH	44	44	-0,108	-
11	11	NAH	NAH	45	45	NAH	NAH
12	12	0,552	Sangat Signifikan	46	46	0,383	Sangat Signifikan
13	13	0,589	Sangat Signifikan	47	47	0,606	Sangat Signifikan
14	14	0,606	Sangat Signifikan	48	48	NAH	NAH
15	15	0,637	Sangat Signifikan	49	49	0,605	Sangat Signifikan
16	16	0,606	Sangat Signifikan	50	50	-0,100	-
17	17	0,597	Sangat Signifikan	51	51	NAH	NAH
18	18	0,313	Sangat Signifikan	52	52	-0,175	-
19	19	NAH	NAH	53	53	NAH	NAH
20	20	NAH	NAH	54	54	NAH	NAH
21	21	NAH	NAH	55	55	0,448	Sangat Signifikan
22	22	0,048	-	56	56	0,606	Sangat Signifikan
23	23	-0,038	-	57	57	0,448	Sangat Signifikan
24	24	0,267	Signifikan	58	58	0,606	Sangat Signifikan
25	25	0,338	Sangat Signifikan	59	59	NAH	NAH
26	26	0,315	Sangat Signifikan	60	60	0,011	-
27	27	0,134	-	61	61	0,011	-
				62	62	0,011	-
				63	63	0,011	-
Catatan: Batas signifikansi koefisien korelasi sebagai berikut:							
df	(N-2)	P=0,05	P=0,01	df	(N-2)	P=0,05	P=0,01
10		0,570	0,700	60		0,250	0,325
15		0,402	0,606	70		0,233	0,302
20		0,423	0,549	80		0,217	0,283
25		0,381	0,496	90		0,205	0,267
30		0,349	0,449	100		0,195	0,254
40		0,304	0,393	125		0,174	0,228
50		0,272	0,354	150		0,159	0,208

Sumber : Hasil Olah Data, 2025

GAMBAR 4. Analisis Validitas

Berdasarkan informasi yang telah dikumpulkan dan analisis yang dilakukan dengan menggunakan ANATES, hasil dari efektivitas pengecoh adalah sebagai berikut:

KUALITAS PENGECHO										KUALITAS PENGECHO																												
*****										*****																												
Jumlah Subyek= 20										Jumlah Subyek= 20																												
Butir Soal= 63										Butir Soal= 63																												
No Butir Baru	No Butir Asli	a	b	c	d	e	*	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
1	1	15**	2-	1+*	1+*	1+*	0	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
2	2	15**	2-	1+*	1+*	1+*	0	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	
3	3	3+*	15**	0+*	2-	0+*	0	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63		
4	4	2-	15**	1+*	1+*	1+*	0	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63			
5	5	5+*	15**	0+*	0+*	0+*	0	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63				
6	6	15**	1+*	4+*	0+*	0+*	0	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63					
7	7	0+*	10**	4+*	0+*	0+*	0	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63						
8	8	0+*	16**	4+*	0+*	0+*	0	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63							
9	9	1+*	0+*	0+*	4+*	15**	0	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63								
10	10	0	20**	0	0	0	0	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63									
11	11	0	0	20**	0	0	0	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63										
12	12	0+*	0+*	0+*	1+*	15**	0	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63											
13	13	14**	0+*	0+*	0+*	0+*	0	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63												
14	14	10**	0+*	0+*	1+*	0+*	0	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63													
15	15	1+*	0+*	0+*	0+*	18**	1+*	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63														
16	16	10**	1+*	0+*	0+*	0+*	0	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63															
17	17	11**	3+*	0+*	6+*	0+*	0	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63																
18	18	0+*	1+*	18**	1+*	0+*	0	50	51	52	53	54	55	56	57	58	59	60	61	62	63																	
19	19	0+*	20**	0	0	0	0	51	52	53	54	55	56	57	58	59	60	61	62	63																		
20	20	20**	0	0	0	0	0	52	53	54	55	56	57	58	59	60	61	62	63																			
21	21	0	0	0	20**	0	0	53	54	55	56	57	58	59	60	61	62	63																				
22	22	0+*	10**	0+*	0+*	1+*	0	54	55	56	57	58	59	60	61	62	63																					
23	23	1+*	0+*	1+*	18**	0+*	0	55	56	57	58	59	60	61	62	63																						
24	24	1+*	1+*	1+*	1+*	16**	0+*	56	57	58	59	60	61	62	63																							
25	25	1+*	1+*	17**	1+*	0+*	0	57	58	59	60	61	62	63																								
26	26	1+*	17**	0+*	2+*	0+*	0	58	59	60	61	62	63																									
27	27	17**	2+*	0+*	0+*	1+*	0	59	60	61	62	63																										

Keterangan:

- ** : Kunci Jawaban
- ++ : Sangat Baik
- + : Baik
- : Kurang Baik
- : Buruk
- : Sangat Buruk

Sumber : Hasil Olah Data, 2025

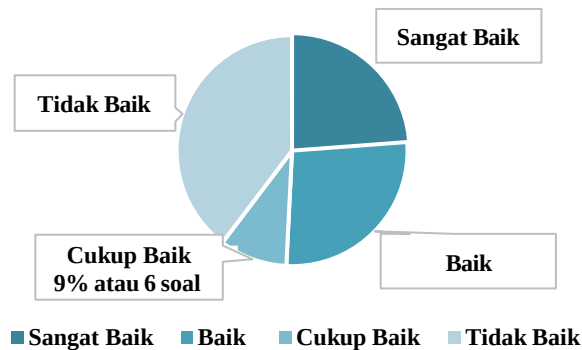
Gambar 4. Kualitas Pengecoh

PEMBAHASAN

Pertama, berdasarkan analisis reliabilitas butir soal yang diberikan ke peserta didik menunjukkan rata-rata skor reliabilitas sebesar 57,30 dengan simpangan baku sebesar 6,33. Pada korelasi antar belahan tes butir soal, angka yang yang diperoleh sebesar 0,94 serta koefisien reliabilitas keseluruhan sebesar 0,97.

Reliabilitas dapat dikatakan baik apabila koefisiennya diatas 0,60. Reliabilitas yang diperoleh pada analisis mencapai 0,97, maka dapat dinyatakan mempunyai tingkat reliabilitas yang sangat tinggi

Berikutnya yang kedua, berdasarkan analisis nilai daya pembeda diklasifikasikan dalam empat kategori, yaitu sangat baik, baik, cukup dan tidak baik (Sabela, O. R., Krisdayanty, D., Taqqiyah, A. Z., Hakim, L., & Pratiwi, V, 2025).

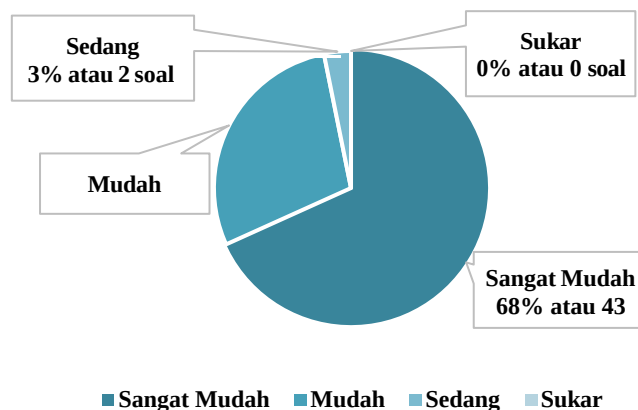


Sumber : Hasil Olah Data, 2025

DIAGRAM 1. Klasifikasi Daya Pembeda

Pengolahan data diperoleh gambaran distribusi daya pembeda sebagaimana ditunjukkan pada diagram. Digram tersebut menunjukkan bahwa 24% atau 15 butir soal berada pada klasifikasi sangat baik, yang berarti mampu membedakan peserta didik secara optimal. Sebanyak 27% atau 17 butir soal termasuk klasifikasi baik, sehingga masih layak digunakan. Selanjutnya 9% atau 6 butir soal terklasifikasi cukup baik sehingga direkomendasikan untuk direvisi. Sementara persentase 40% atau 25 butir soal berada klasifikasi tidak baik yang menunjukkan bahwa butir soal tersebut memerlukan revisi secara keseluruhan.

Selanjutnya yang ketiga, hasil evaluasi tingkat kesukaran terhadap butir soal menggunakan perangkat lunak Anates dengan aspek tingkat kesulitan dikelompokkan dalam tiga kategori, yaitu sulit, sedang, dan mudah (Ryzca Siti Qomariyah dkk., 2022).



Sumber : Hasil Olah Data, 2025

DIAGRAM 2. Klasifikasi Tingkat Kesukaran

Hasil olah data diperoleh bahwa 68% atau 43 butir soal masuk kelompok sangat mudah, kemudian 29% atau 18 butir soal berada kelompok mudah. Selanjutnya, hanya 3% atau 2 butir soal yang termasuk kelompok sedang dan tidak ada soal yang termasuk kelompok sukar. Hal ini menjadi indikasi bahwa alat ukur tes tersebut cenderung tidak memberikan tantangan yang cukup dan belum mampu membedakan kemampuan siswa secara efektif.

Butir soal yang tergolong mudah dan sangat mudah dapat meningkatkan rasa percaya diri siswa, namun jika jumlahnya terlalu banyak, bisa menyebabkan penilaian tidak mencerminkan kemampuan sebenar siswa. Di sisi lain, hanya sedikit soal yang berada dalam kategori sedang, sehingga variasi tingkat kesulitan dalam alat ukur ini masih tidak seimbang. Untuk evaluasi di masa mendatang, disarankan agar butir soal dengan kelompok sangat mudah direvisi atau diganti dengan buti soal memiliki tingkat kesulitan yang lebih tinggi. Tujuannya adalah untuk menciptakan distribusi tingkat kesulitan yang lebih seimbang dan tes dapat mengukur kemampuan peserta didik dengan lebih menyeluruh.

Sedangkan, berdasarkan analisis, teridentifikasi 32 butir soal dalam klasifikasi valid, sementara 31 butir soal lainnya dalam klasifikasi tidak valid. Tabel berikut menyajikan hasil analisis butir soal terkait klasifikasi validitas.

TABEL 5. Klasifikasi Validitas

Kategori	Jumlah	Butir Soal
Valid	32	1 - 9, 12 - 18, 24 - 26, 32, 33, 35, 36, 38, 39, 46, 47, 49, 55 - 58
Tidak Valid	31	10, 11, 19 - 23, 27 - 31, 34, 37, 40 -45, 48, 50 - 54, 59 - 63

Sumber : Hasil Olah Data, 2025

Menurut analisis yang didapat, sebanyak 63 butir soal, diantaranya 32 butir soal dinyatakan valid, yaitu soal yang terletak di nomor 1, 2, 3, 4, 5, 6, 7, 8, 9, 12, 13, 14, 15, 16, 17, 18, 24, 25, 26, 32, 33, 35, 36, 38, 39, 46, 47, 49, 55, 56, 57, dan 58. Sedangkan 31 butir soal lainnya, yang mencakup nomor 10, 11, 19, 20, 21, 22, 23, 27, 28, 29, 30, 31, 34, 37, 40, 41, 42, 43, 44, 45, 48, 50, 51, 52, 53, 54, 59, 60, 61, 62, serta 63, dianggap tidak valid. Dari hasil tersebut, dapat disimpulkan bahwa tingkat validitas berada pada kategori yang cukup, di mana hampir setengah dari jumlah butir soal dinyatakan layak digunakan. Butir soal yang valid sebaiknya digunakan untuk evaluasi di masa yang akan datang, sedangkan butir soal yang tidak valid memerlukan perbaikan lebih lanjut.

Berdasarkan analisis menggunakan ANATES, kualitas pengecoh terbagi menjadi lima aspek, yaitu sangat baik, baik, kurang baik, buruk, dan sangat buruk.

TABEL 6. Klasifikasi Validitas

Kategori	Jumlah	Butir Soal
Sangat Baik	46	1 - 7, 17 - 19, 21 - 27, 33 - 36, 38, 39, 41 - 63
Baik	10	8 - 16, 20
Kurang Baik	8	2 - 7, 12, 13
Buruk	12	2 - 6, 8 - 14
Sangat Buruk	6	8,9 13 - 16

Sumber : Hasil Olah Data, 2025

Hasil analisis untuk mengevaluasi kualitas pengecoh, ditemukan 46 butir soal yang memiliki pengecoh beraspek sangat baik, yaitu pada nomor 31, 2, 3, 4, 5, 6, 7, 17, 18, 19, 21, 22, 23, 24, 25, 26, 27, 33, 34, 35, 36, 38, 39, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53,

54, 55, 56, 57, 58, 59, 60, 61, 62, dan 63. Pengecoh di butir-butir tersebut berfungsi dengan efektif sekali karena berhasil menarik perhatian siswa yang masih kurang memahami materi, sehingga berkontribusi dalam memisahkan peserta didik yang memahami dan yang belum sepenuhnya menguasai materi dengan baik. Selanjutnya, ada 10 butir soal dengan pengecoh yang tergolong baik, yakni pada nomor 8, 9, 10, 11, 12, 13, 14, 15, 16, dan 20. Hal ini mengindikasikan bahwa pengecoh dalam butir tersebut masih cukup efektif meski daya tariknya terhadap peserta yang salah menjawab tidak sekuat yang berada dalam kategori sangat baik. Terakhir, ada 8 butir soal memiliki pengecoh yang tergolong kurang baik, yaitu pada nomor 2, 3, 4, 5, 6, 7, 12, dan 13. Pengecoh di kategori ini masih dapat digunakan, akan tetapi memerlukan perbaikan agar lebih menarik bagi siswa yang belum mengerti konsep di dalam soal.

Sementara itu, ada 12 butir soal yang memiliki pengecoh dengan kategori buruk, yakni di nomor 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, dan 14. Distraktor pada soal-soal ini tidak berfungsi dengan optimal karena hanya sedikit, atau bahkan tidak ada, peserta yang memilihnya. Sementara itu, terdapat 6 butir soal yang memiliki pengecoh dengan kategori sangat buruk, yakni pada nomor 8, 9, 13, 14, 15, dan 16. Ini menunjukkan bahwa pengecoh untuk butir tersebut sama sekali tidak berfungsi, sehingga perlu adanya perbaikan pada pilihan jawaban untuk meningkatkan kualitas butir soal.

Pengecoh dinyatakan sangat efisien apabila mampu menarik perhatian para responden. Menyusun pilihan pengecoh dalam suatu penilaian bukan hal sederhana dikarenakan berpengaruh langsung terhadap kemampuan butir soal dalam membedakan tingkat penguasaan peserta didik. Apabila pengecoh disusun dengan kurang tepat, kemampuan butir soal untuk membedakan peserta didik yang menguasai dan kurang menguasai materi akan menurun. Sedangkan, pengecoh yang baik dapat meningkatkan kualitas butir soal serta menghasilkan evaluasi lebih akurat mengenai kemampuan peserta didik.

SIMPULAN

Penelitian ini menyimpulkan bahwa meski soal pilihan ganda dalam Dokumen Berbasis Digital sangat dipercaya (0,97), kualitas soalnya secara keseluruhan belum sesuai harapan. Sebagian besar soal (97%) terlalu mudah, hampir 40% tidak mampu membedakan kemampuan siswa, dan hampir separuh dari seluruh soal dianggap tidak valid. Temuan ini menunjukkan meskipun instrumen terpercaya, kualitas soal dari segi tingkat kesulitan, kemampuan membedakan, dan efektivitas pengecoh masih kurang memadai.

Metodologi penelitian ini adalah deskriptif kuantitatif, dengan fokus pada analisis soal menggunakan Teori Tes Klasik. Namun, belum ada penelitian mendalam mengenai aspek kualitatif soal (seperti keabsahan isi dan ulasan para ahli), serta faktor-faktor lain (seperti metode pengajaran atau kemampuan awal siswa) yang menyebabkan banyak soal mudah dan kemampuan pembedanya rendah.

Untuk penelitian ke depan, disarankan agar analisis kuantitatif soal dipadukan dengan metode kualitatif, seperti wawancara dengan guru atau siswa, untuk menemukan penyebab utama masalah. Penelitian berikutnya juga perlu menitikberatkan pada pengembangan dan pengujian instrumen yang telah diperbaiki, serta melakukan studi perbandingan menggunakan Teori Respons Butir (IRT) guna memperoleh pemahaman lebih mendalam.

DAFTAR PUSTAKA

1. Agustin, Livia, Gina Zahra Wildaniar Amraow, dan Luqman Hakim. (2024). "Analisis Butir Soal Pilihan Ganda pada Elemen Akuntansi Lembaga atau Instansi Pemerintah Menggunakan Program Komputer Anates."

2. Ainun, Andi Ainun Lady R., Ramly, dan Jafar Ahiri. (2024). "Analisis Kualitas Tes Buatan Guru Mata Pelajaran Akuntansi Keuangan." *Accounting: Jurnal Pendidikan Akuntansi* 4(3):136–47. doi:10.36709/jpa.v3i1.59.
3. Alpusari, Mahmud. (2015). "Analisis Butir Soal Konsep Dasar Ipa 1 Melalui Penggunaan Program Komputer Anates Versi 4.0 For Windows." *Primary: Jurnal Pendidikan Guru Sekolah Dasar* 3(2):106. doi:10.33578/jpkip.v3i2.2501.
4. Amelia, Nurul, dan Ibnu Sina. (2021). "Analisis butir soal penilaian akhir tahun mata pelajaran matematika menggunakan software anates." 12(3).
5. Anshari, Muhammad Isa, Rodiah Nasution, Muhammad Irsyad, Alifia Zuhriatul Alifa, dan Indah Aminatus Zuhriyah. 2024. "Analisis Validitas dan Reliabilitas Butir Soal Sumatif Akhir Semester Ganjil Mata Pelajaran PAI." *EDUKATIF: JURNAL ILMU PENDIDIKAN* 6(1):964–75. doi:10.31004/edukatif.v6i1.5931.
6. Asria, Lailatul. (2023). "Analisis Butir Soal Penilaian Tengah Semester (PTS) Matematika Kelas XI Berdasarkan Teori Klasik." 4(1).
7. Asyarah, Serlianti Putri, Arina Candra Febriyanti, dan Cheri J. C. Sitanggang. (2025). "Analisis Butir Soal Hots Pada Elemen Perpajakan Fase F SMK Akuntansi Menggunakan Software Anates."
8. Augustia, Amanda Diskafri, Charisma Nurochmatiana Agustia, Divanida Azzahra, Luqman Hakim, dan Vivi Pratiwi. (2025). "Analisis Validitas dan Reliabilitas Soal Pilihan Ganda dengan Menggunakan Software Anates pada Mata Pelajaran Perpajakan." *AKADEMIK: Jurnal Mahasiswa Ekonomi & Bisnis* 5(1):250–65. doi:10.37481/jmeh.v5i1.1165.
9. Cahyani, Deva Arinda, Annisa Hanum Aurelia, Luqman Hakim, dan Vivi Pratiwi. (2024). "Analisis Butir Soal HOTS Akuntansi Perusahaan Jasa Menggunakan Software Anates pada Siswa SMK Kelas 1."
10. Firandhika, Ahmad, Kartika Dwi Apriliyanti, Andini Kasih Agus Saputri, dan Vivi Pratiwi. (2024). "Analisis Butir Soal HOTS Pilihan Ganda Pada Elemen Komputer Akuntansi Di SMKN 10 Surabaya Menggunakan Aplikasi Anates."
11. Fitraynsyah, Muhammad Andryan, dan Fitri Hilmiyati. (2024). "Peran Tingkat Kesukaran dan Daya Pembeda dalam Analisis Butir Tes: Kajian Literatur untuk Pendidikan Menengah." 1(4).
12. Handayani, Sri. (2022). "Analisis uas biolog(Purwati dkk. 2021)i kelas x dengan teori tes klasik dan item response theory (rasch model)." *BIO-PEDAGOGI* 11(2):76. doi:10.20961/bio-pedagogi.v11i2.63177.
13. Hari, Hari Purnomo Susanto, Siti Irene Astuti D, dan Endang Mulyani. (2023). "Penskoran yang Fair pada Tes Matematika Berbentuk Pilihan Ganda Menggunakan Item Response Theory." *JMPM: Jurnal Matematika dan Pendidikan Matematika* 8(2):157–72. doi:10.26594/jmpm.v8i2.3293.
14. Hartono, Inayah Dzil Izzati, Andi Besse Tenriawaru, dan Kurnia Ningsih. (2024). "Analisis Butir Soal Penilaian Sumatif Ipa Kelas Vii Smp Negeri 3 Pontianak Menggunakan Anates." *Jurnal Kajian Pembelajaran dan Keilmuan* 8(2):162–71. doi:10.26418/jurnalkpk.v8i2.78282.

15. Ikhsaniyah, Salsabella Nur, Amanda Dwi Kurnia, Maharani Zuroida, Vivi Pratiwi, dan Luqman Hakim. (2024). "Analisis Butir Soal Perpajakan Pph Pasal 21 Menggunakan Software Anates Pada Pendekatan Teori Tes Klasik." 12(2).
16. Iriyanti, Rosita Dewi, Najwa Dafa Agustina, Dini Aulia Prinita, dan Vivi Pratiwi. (2024). "Analisis Butir Soal Pilihan Ganda Mata Pelajaran Ekonomi Bisnis dan Administrasi Umum dengan Menggunakan Aplikasi Software Anates di SMK NU Gresik." *Jurnal Pengabdian Kepada Masyarakat* 04(02).
17. Kafi, Fina Aunul, dan Zidni Ilma. (2022). "Analisis Soal Tematik Bahasa Arab berbasis HOTS pada Kelas XII MAN 3 Jember." *Al-Fusha : Arabic Language Education Journal* 4(1):1-9. doi:10.62097/alfusha.v4i1.729.
18. Karindi, Dito Risky dan Rufi'i. (2024). "Penggunaan Anates 4.0.9 dalam menguji kelayakan soal pilihan ganda sumatif matematika." *LINEAR: Journal of Mathematics Education* 5(2):181-96. doi:10.32332/2zg4f753.
20. Kumalasari, Ellisia, Jamilah Karaman, Dyah Mustikasari, dan Fajar Kurniawan. (2022). "Analisis Butir Soal Pada Tes Seleksi Perangkat Desa Berbasis Computer Assisted Test (Cat) Sebagai Bentuk Proses Evaluasi." . . *ISSN* 7.
21. Masruri, Faizah Balita, Firda Nur Farila, Michaelia Alriviany Putri, dan Vivi Pratiwi. (2024). "Artikel Optimasi Evaluasi Pembelajaran: Analisis Butir Soal Pilihan Ganda dengan Aplikasi Anates." 8.
22. Maulana, Andi. (2022). "Analisis Validitas, Reliabilitas, dan Kelayakan Instrumen Penilaian Rasa Percaya Diri Siswa." *Jurnal Kualita Pendidikan* 3(3):133-39. doi:10.51651/jkp.v3i3.331.
23. Mulyati, Budi. (2021). "Analisis Butir Soal Uraian Mata Kuliah Pengantar Akuntansi 1." *Progress: Jurnal Pendidikan, Akuntansi dan Keuangan* 4(2):178-86. doi:10.47080/progress.v4i2.1469.
24. Mustaqim, Mustaqim, dan Hidayu Sulisti. (2024). "Analisis Butir Soal Pas Matematika Peminatan: Daya Pembeda, Tingkat Kesukaran, Dan Kualitas Pengecoh." *Al-Adad : Jurnal Tadris Matematika* 3(1):44-56. doi:10.24260/add.v3i1.3011.
25. Pradita, Eliza, dan Priarti Megawanti. (2023). "Analisis Tingkat Kesukaran, Daya Pembeda, dan Fungsi Distraktor PTS Matematika SMPN Jakarta."
26. Purwati, Ledi Merlin, Rizky Arianty, Diah Malaka Syakilah, Saiful Ridlo, dan Endang Susilaningsih. (2021). "Analisis Soal Tes Pilihan Ganda Berbasis Higher Order Thinking Skill menggunakan Aplikasi Anates Windows Versi 4.0.9 For Windows." *Jurnal Pendidikan UNIGA* 15(2):460. doi:10.52434/jp.v15i2.1287.
27. Puteri, Nasywa Dellia, Nazah Dwi Putricia, Khoirina Arsyanti, Nazah Dwi Putricia, Luqman Hakim, dan Vivi Pratiwi. (2025). "Analisis Kelayakan Butir Soal Pada Handout Akuntansi Keuangan Dana Kas Kecil Berbasis E-Book Menggunakan Software Anates." *JIBEMA: Jurnal Ilmu Bisnis, Ekonomi, Manajemen, dan Akuntansi* 2(3):205-14. doi:10.62421/jibema.v2i3.108.
28. Qomariyah, Ryzca Siti. (2022). "Analisis Tingkat Kesukaran Dan Daya Pembeda Pada Butir Soal Pilihan Ganda Mata Pelajaran Bahasa Indonesia Kelas V Semester 1 SDN Kedungdalem 2." 1(2).

29. Sabela, Olivia Rika, Dewi Krisdayanty, Az Zahra Taqqiyah, Lukman Hakim, dan Vivi
30. Pratiwi. (2025). "Analisis Butir Soal HOTS Elemen Dokumen Berbasis Digital (FASE E) Menggunakan Program Anates." *Education Achievement: Journal of Science and Research* 251–62. doi:10.51178/jsr.v6i1.2328.

PROFIL SINGKAT

Betty Sofi Hariati adalah Mahasiswa Program Studi Pendidikan Akuntansi, Fakultas Ekonomika dan Bisnis Universitas Negeri Surabaya.

Luqman Hakim adalah Dosen Program Studi Pendidikan Akuntansi, Fakultas Ekonomika dan Bisnis Universitas Negeri Surabaya.

Vivi Pratiwi adalah Dosen Program Studi Pendidikan Akuntansi, Fakultas Ekonomika dan Bisnis Universitas Negeri Surabaya.

Ainuz Zakia adalah Mahasiswa Program Studi Pendidikan Akuntansi, Fakultas Ekonomika dan Bisnis Universitas Negeri Surabaya.

Revalina Cintia Ayu adalah Mahasiswa Program Studi Pendidikan Akuntansi, Fakultas Ekonomika dan Bisnis Universitas Negeri Surabaya.

Andrea Dzakira Maheswari adalah Mahasiswa Program Studi Pendidikan Akuntansi, Fakultas Ekonomika dan Bisnis Universitas Negeri Surabaya.

Shafira Eka Salsabila adalah Mahasiswa Program Studi Pendidikan Akuntansi, Fakultas Ekonomika dan Bisnis Universitas Negeri Surabaya.

Qiptiyah Qurrotul Aini adalah Mahasiswa Program Studi Pendidikan Akuntansi, Fakultas Ekonomika dan Bisnis Universitas Negeri Surabaya.

Tri Dwi Rahmawati adalah Mahasiswa Program Studi Pendidikan Akuntansi, Fakultas Ekonomika dan Bisnis Universitas Negeri Surabaya.