



Pemanfaatan Software Anates Sebagai Evaluasi Butir Soal pada Mata Pelajaran Ekonomi Bisnis dan Administrasi Umum Fase F

Latifa Nurul Hidayati^{1*}, Luqman Hakim², Vivi Pratiwi³, Selvi Agustin⁴, Nabila Aulia Sari⁵, Imro Atus Sholehah⁶, Alfi Rizka Chafidho Hanum⁷, Sarah Fidela Rahayu⁸

¹⁻⁸ Universitas Negeri Surabaya, Indonesia

*Penulis Korespondensi: 24080304060@mhs.unesa.ac.id

Abstract. *This study was conducted to review the quality of test items in Business Economics and General Administration subjects using a classical analysis approach. The assessment process included examining validity, reliability, difficulty level, the ability of questions to differentiate student abilities, and the accuracy of distractors. The research approach used was descriptive quantitative, with 32 students from grades XI and XII majoring in Accounting as subjects and 50 multiple-choice questions as instruments. The main focus of this study was to ensure that the evaluation tools used were truly capable of describing students' competency achievements. Data analysis was conducted through observation results and student responses on Google Forms, which were then processed using the Anates Version 4.0 application. The results of the analysis showed that most of the questions had high validity and reliability, good discrimination power, and a moderate level of difficulty. However, the overall quality of the distractors was not good enough, and there were no questions in the difficult category. Overall, the quality of the questions in this subject can be categorized as good and suitable for evaluating the learning of vocational high school students in grades XI and XII.*

Keywords: *Anates; Business Economic; General Administration; HOTS Questions; Item Analysis.*

Abstrak. Penelitian ini dilakukan untuk meninjau kembali kualitas butir soal pada mata pelajaran Ekonomi Bisnis dan Administrasi Umum dengan memanfaatkan pendekatan analisis klasik. Proses penilaian meliputi pemeriksaan validitas, reabilitas, tingkat kesukaran, kemampuan soal dalam membedakan kemampuan siswa, serta ketepatan fungsi pengecoh. Pendekatan penelitian yang digunakan adalah deskriptif kuantitatif, dengan subjek sebanyak 32 siswa SMK kelas XI dan XII jurusan Akuntansi serta instrumen berupa 50 butir soal pilihan ganda. Fokus utama studi ini adalah memastikan bahwa alat evaluasi yang digunakan benar – benar mampu menggambarkan pencapaian kompetensi siswa. Analisis data dilakukan melalui hasil observasi dan respon siswa pada Google Form yang kemudian diolah menggunakan aplikasi Anates Versi 4.0. Hasil analisis sebagian besar soal menunjukkan validitas dan reabilitas yang tinggi, daya pembeda yang tergolong baik, dan tingkat kesukaran yang berada pada kategori sedang. Namun, kualitas pengecoh secara keseluruhan belum cukup baik dan belum terdapat soal dengan kategori sulit. Secara keseluruhan, kualitas butir soal pada mata pelajaran ini dapat dikategorikan baik dan layak digunakan untuk evaluasi pembelajaran siswa SMK kelas XI dan XII.

Kata Kunci: Administrasi Umum; Analisis Butir Soal; Anates; Butir Soal HOTS; Ekonomi Bisnis.

1. LATAR BELAKANG

Pendidikan kejuruan pada program Akuntansi di SMK dituntut menghasilkan lulusan yang siap bekerja, memiliki kompetensi teknis yang memadai, serta sikap profesional (Anazar & Munir, 2024). Mata pelajaran Ekonomi Bisnis dan Administrasi Umum pada Fase F Kurikulum Merdeka berperan penting sebagai dasar dalam mengembangkan kompetensi kewirausahaan, kemampuan mengelola administrasi, serta pemahaman tentang proses bisnis yang relevan bagi siswa SMK Akuntansi. Namun dalam kegiatan pembelajaran, kualitas soal-soal yang digunakan oleh guru masih bervariasi. Banyak soal yang belum dianalisis secara kuantitatif sehingga tidak bisa dipastikan validitas, reliabilitas, tingkat kesukaran, dan daya pembeda soal belum terjamin (Ikawati et al., 2024). Hal ini bisa membuat penilaian tidak akurat

dalam mengetahui siswa, dan berpengaruh pada pengambilan keputusan dalam pembelajaran serta perbaikan materi pembelajaran. Beberapa penelitian menunjukkan masih banyak soal yang belum memenuhi standar jika dianalisis menggunakan perangkat lunak seperti anates.

Analisis soal-soal menggunakan perangkat lunak semakin banyak digunakan di lingkungan sekolah menengah SMP, SMA, dan SMK. Anates adalah salah satu aplikasi yang sering digunakan untuk menguji kualitas soal, sehingga menghasilkan berbagai statistik seperti validitas, reliabilitas, indeks kesukaran, daya pembeda, dan efektivitas pengecoh (Anike, 2022). Dengan begitu, guru bisa memperbaiki instrumen soal berdasarkan data. Penelitian terkini menemukan tiga hal penting. Pertama, Anates bisa mempercepat dan membuat proses analisis soal lebih rapi. Kedua, penggunaan anates membantu meningkatkan kualitas soal dalam berbagai penelitian. Ketiga, penggunaan anates di sekolah masih terbatas, sering hanya dipakai untuk urusan administratif tanpa terhubung dengan proses perbaikan pembelajaran secara teratur. Hasil ini menunjukkan ada kesempatan untuk menggunakan anates lebih terorganisir lagi dalam konteks kurikulum merdeka fase F (Damayanti et al., 2024).

Penelitian ini didasarkan pada hubungan antara tiga komponen penting dalam proses evaluasi pembelajaran. Pertama, penyusunan instrumen evaluasi, yaitu perancangan soal pada mata pelajaran Ekonomi Bisnis dan Administrasi Umum. Kedua, analisis kualitas butir soal, yang mencakup uji validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh menggunakan software Anates. Ketiga, tindak lanjut instruksional, yaitu perbaikan atau revisi soal serta penyesuaian strategi pembelajaran setelah hasil analisis diperoleh. Penggunaan Anates mendukung prinsip asesmen autentik yang ditekankan dalam Kurikulum Merdeka, yaitu asesmen yang memberikan informasi yang akurat untuk memperbaiki proses pembelajaran dan membantu pembentukan Profil Pelajar Pancasila (Andriatoni., Adia, B., Yessari, M., Yeni, 2025). Oleh karena itu, analisis butir soal menggunakan perangkat seperti Anates menjadi sangat relevan dalam konteks kebijakan pendidikan saat ini.

Penelitian ini penting karena memberikan manfaat praktis dan teoritis. Secara praktis, penelitian membantu guru SMK Akuntansi menghasilkan soal yang lebih valid dan reliabel sehingga penilaian lebih akurat. Penelitian ini juga mendukung implementasi Fase F Kurikulum Merdeka melalui asesmen berbasis data. Secara teoritis, penelitian mengisi kekosongan kajian terkait penggunaan Anates pada mata pelajaran Ekonomi Bisnis dan Administrasi Umum di SMK, serta memberikan rekomendasi tentang integrasi analisis butir otomatis dalam evaluasi pembelajaran. Temuan penelitian sebelumnya juga menunjukkan bahwa penggunaan Anates dapat meningkatkan kualitas penilaian di tingkat sekolah menengah (Adimas et al., 2025).

2. KAJIAN TEORITIS

Pentingnya Evaluasi Dalam Pembelajaran

Evaluasi adalah kegiatan yang dilakukan oleh pendidik guna mengetahui, memahami, dan memanfaatkan hasil pembelajaran peserta didik dalam mencapai sasaran yang telah ditentukan. Proses evaluasi seharusnya berlandaskan pada periode dan waktu, bukan hanya sesaat. Ini menunjukkan bahwa evaluasi adalah hasil akhir dari serangkaian pengukuran yang dilakukan berulang kali dengan tujuan tertentu (Purnomo et al., 2022). Hal yang penting dari evaluasi program meliputi penyediaan informasi yang dipakai sebagai dasar mengambil kebijakan dan keputusan, mengukur hasil yang diraih oleh peserta didik, menilai kurikulum, serta melakukan perbaikan terhadap materi dan program pendidikan.

Evaluasi dalam pembelajaran memiliki peranan yang sangat krusial untuk mengevaluasi sejauh mana tujuan pembelajaran telah dicapai. Bukan hanya sekadar alat untuk menguji pengetahuan peserta didik, namun juga untuk mengidentifikasi kekuatan dan kelemahan dalam pengajaran yang digunakan oleh pendidik. Penilaian tersebut dapat memberikan gambaran jelas tentang apa yang telah dipahami peserta didik dan bagian mana yang membutuhkan perhatian lebih. Dan sebagai peserta didik pun dapat mengetahui dimana posisi mereka dalam proses pembelajaran, sehingga mereka dapat memperbaiki pemahaman terhadap materi yang belum dikuasai.

Analisis Butir Soal

Kegiatan menganalisis butir soal merupakan langkah yang perlu dilakukan oleh seorang pendidik dengan tujuan untuk mengidentifikasi soal-soal yang baik, yang kurang baik, serta soal yang berkualitas rendah. Analisis mengenai kualitas tes dan butir soal melibatkan berbagai langkah yang dilaksanakan untuk memastikan bahwa tes dan butir soal dapat diverifikasi dan dapat digunakan secara efektif. Penilaian terhadap kualitas tes dan butir soal merupakan langkah penting dalam mengembangkan serta mengevaluasi instrumen dalam penelitian. Kualitas yang tinggi dari tes dan masing-masing pertanyaan menjamin bahwa tes tersebut mampu menghasilkan data yang sah, dapat diandalkan, dan sesuai dengan kemampuan atau variabel yang ingin diukur. Melalui analisis menyeluruh terhadap setiap pertanyaan, kita dapat menemukan kelemahan yang ada dan melakukan perbaikan yang diperlukan. Langkah ini sangat penting untuk memastikan bahwa pertanyaan yang diajukan menghasilkan hasil yang sah, dapat diandalkan, dan sesuai untuk peserta didik (Artama et al., 2023).

HOTS (*High Order Thinking Skill*)

HOTS adalah kemampuan atau cara berpikir seseorang pada tingkat yang lebih tinggi. Ada beberapa istilah yang berhubungan dengan kemampuan berpikir, yang secara nyata

memiliki perbedaan yaitu berpikir tingkat tinggi, berpikir kompleks, dan berpikir kritis. Bloom menjelaskan tingkatan dari proses kognitif mulai dari yang paling dasar hingga yang lebih rumit, yang dikenal dengan istilah level keterampilan kognitif. Tingkatan tersebut terdiri dari enam tingkat yaitu *knowledge*, *understand*, *application*, *analysis*, *synthesis*, dan *evaluation*. Dimensi proses berpikir dalam Taksonomi Bloom yang telah diperbarui dan disempurnakan oleh Anderson & Krathwohl, terdiri dari C1-*knowing* (mengetahui), C2-*understanding* (memahami), C3-*aplying* (menerapkan), C4-*analyzing* (menganalisis), C5-*evaluating* (mengevaluasi), C6-*creating* (menciptakan).

HOTS dapat dijelaskan sebagai kemampuan yang memerlukan pemikiran kritis dan kreatif untuk menyelesaikan masalah. Seseorang bisa menganalisis, menghubungkan, merinci, serta memahami masalah untuk menemukan solusi atau gagasan baru. Pertanyaan-pertanyaan HOTS biasanya menilai kemampuan dalam kategori C4 hingga C6, yaitu menganalisis, mengevaluasi, dan mengkreasi (Manik et al., 2020).

Pemanfaatan Anates versi 4.0 dalam Analisis Butir Soal

Software Anates Versi 4.0. banyak dimanfaatkan oleh para pendidik dan peneliti di bidang pendidikan untuk melakukan analisis kualitas butir soal secara statistik. Analisis mengenai validitas, reabilitas, tingkat kesulitan, daya pembeda, kualitas pengecoh, serta rekapan dari analisis butir soal merupakan fitur yang disediakan pada perangkat ini. Keunggulan paling signifikan dari alat ini adalah antar muka yang mudah digunakan oleh metode analisis yang efisien dan tepat (Purwati et al., 2021). Hasil tes tersebut dapat dimanfaatkan untuk memperbaiki soal yang masih belum optimal sebelum digunakan sekali lagi. Dengan demikian, kualitas penilaian akan meningkat, dan siswa akan mendapatkan kesempatan untuk dinilai dengan cara yang adil dan akurat (Jatmoko et al., 2025). Penggunaan Anates sangat mendukung guru dalam melakukan evaluasi butir soal dan mendorong dalam pengambilan keputusan yang berbasis data saat menyusun kembali soal untuk penilaian selanjutnya.

3. METODE PENELITIAN

Penelitian ini adalah suatu evaluasi yang menggunakan pendekatan kuantitatif deskriptif. Pendekatan ini dipilih karena penelitian ini tidak menguji hipotesis, namun lebih menekankan pada deskripsi kondisi nyata pada instrumen penilaian melalui data yang diperoleh dari respon siswa. Data responden yang berbentuk nilai numerik diolah dengan perangkat lunak Anates V4. Hasil pengolahan tersebut kemudian diinterpretasikan secara deskriptif melalui persentase. Penelitian ini bertujuan untuk menilai kualitas butir soal berbasis

HOTS yang telah dikembangkan dan diberikan kepada siswa kelas XI dan XII jurusan Akuntansi. Evaluasi mencakup aspek validitas, reliabilitas, tingkat kesulitan, daya pembeda dan kualitas pengecoh menggunakan kuesioner yang terdiri dari 50 soal pilihan ganda.

Subjek penelitian terdiri dari siswa kelas XI dan XII SMK dari jurusan Akuntansi. Sampel diperoleh melalui teknik random sampling, menghasilkan 32 responden yang berpartisipasi secara sukarela. Instrumen yang digunakan berupa tes objektif pilihan ganda dengan 50 butir soal yang terdiri dari lima opsi jawaban (A hingga E). Tes tersebut dirancang dengan format evaluasi pilihan ganda yang memungkinkan siswa untuk memilih jawaban yang dianggap paling tepat (Fietri et al., 2021).

Pengumpulan data dilakukan melalui kuesioner yang umum digunakan yakni Google Form dan disebarkan kepada para siswa kelas XI dan XII jurusan Akuntansi. Materi tes mencakup topik Ekonomi Bisnis dan Administrasi Umum yang mencakup kisi-kisi, silabus, serta capaian pembelajaran. Data yang terkumpul kemudian dievaluasi menggunakan perangkat lunak Anates V4.0. (Fiska et al., 2021). Analisis mencakup validitas setiap butir soal yang menggunakan korelasi Product Moment, serta reliabilitas yang diukur dengan Cronbach Alpha dan dianggap baik bila nilai $\alpha > 0,70$. Selain itu, analisis tingkat kesukaran soal ini untuk menggolongkan butir soal ke dalam tiga kategori, yaitu mudah, sedang, dan sukar. Di samping itu, daya pembeda soal digunakan untuk menilai sejauh mana setiap butir soal dalam membedakan siswa yang mampu menjawab dan kurang bisa menjawab. Penilaian mutu butir soal mengharuskan dilakukannya analisis mendalam terhadap empat kriteria pokok, yakni reliabilitas, validitas, daya pembeda, serta tingkat kesukaran. Hasil pengujian disajikan dalam bentuk narasi deskriptif dengan mengintegrasikan data numerik yang dihasilkan dari perhitungan menggunakan perangkat lunak Anates Versi 4.0.

4. HASIL DAN PEMBAHASAN

Analisis butir soal adalah langkah penting yang dilakukan pendidik untuk mengevaluasi kualitas soal dalam suatu tes. Proses ini mencakup beberapa tahapan yang saling berkaitan, yaitu pembuatan soal, penggunaan soal dalam evaluasi pembelajaran, perbaikan soal berdasarkan hasil analisis, dan penilaian kualitas soal secara menyeluruh (Firandhika et al., 2024). Kegiatan analisis ini merupakan upaya untuk memaksimalkan mutu soal yang akan diberikan kepada para siswa.

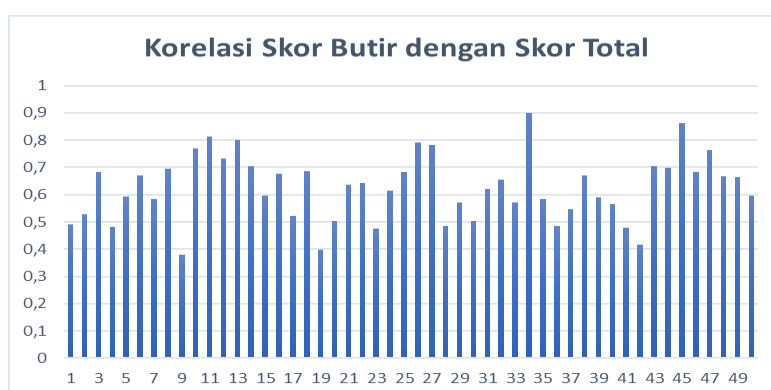
Sesuai dengan hasil penelitian yang sudah dilaksanakan yaitu dengan menguji hasil tes peserta didik tingkat SMK Jurusan Akuntansi. Penelitian ini mempunyai tujuan untuk menilai kualitas butir soal HOTS (*Higher Order Thinking Skills*) dari yang mudah, sedang, hingga

sukar berbentuk pilihan ganda yang berjumlah 50 butir soal pada mata pelajaran Ekonomi Bisnis dan Administrasi Umum Fase F. Analisis ini diharapkan dapat memberikan gambaran mengenai soal-soal yang berkualitas dan sesuai dengan tingkat pemahaman peserta didik.

Validitas

Validitas adalah ukuran yang krusial untuk memastikan bahwa pertanyaan yang dirancang sudah relevan dengan sasaran pembelajaran (Hakim, 2025). Analisis validitas pada Anates bertujuan untuk mengevaluasi seberapa baik soal-soal dalam tes mampu mengukur hal yang seharusnya diukur, yaitu kompetensi peserta didik terhadap materi yang sudah diajarkan.

Dalam Anates, analisis validitas dilakukan dengan cara menghitung indeks validitas untuk setiap pertanyaan. Hasil dari analisis tersebut membantu pendidik dalam memutuskan apakah pertanyaan tersebut sudah sesuai atau perlu diperbaiki agar lebih akurat dalam mengukur kemampuan peserta didik. Hasil analisis dan hasil uji validitas ditunjukkan pada gambar 1.



Keterangan: Hasil Korelasi Skor Butir Soal.

Sumber: data diolah (2025).

Gambar 1. Hasil Korelasi Skor Butir Soal.

Tabel 1. Validitas.

Koefisien	Kategori
0,80 - 1,00	Validasi soal sangat tinggi
0,60 - 0,80	Validasi soal tinggi
0,40 - 0,60	Validasi soal cukup
0,20 - 0,40	Validasi soal rendah
0,00 - 0,20	Validasi soal sangat rendah

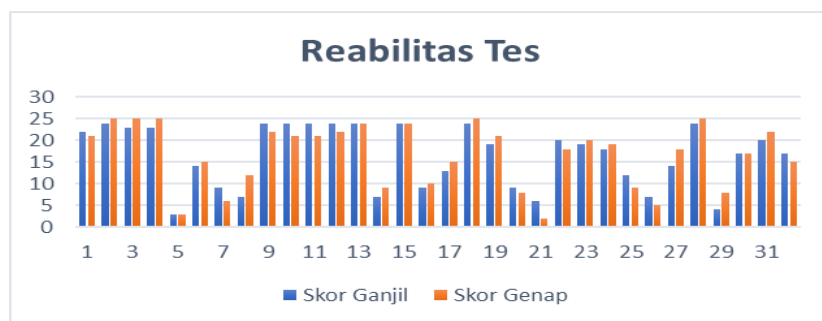
Sumber : Data diolah (2025).

Analisis validitas instrumen tes menunjukkan bahwa dari keseluruhan 50 butir soal pilihan ganda yang diuji memiliki validitas 100% dengan kategori "sangat signifikan". Seluruh

butir soal menunjukkan nilai koefisien validitas (r -hitung) yang melampaui nilai r -tabel, sehingga seluruh instrumen tes dinyatakan valid. Temuan ini membuktikan bahwa setiap butir soal telah berhasil mengukur aspek kompetensi atau pengetahuan siswa SMK pada mata pelajaran ini sudah sesuai dengan tujuan kurikulum. Tingginya validitas butir soal memberikan jaminan bahwa data hasil tes dapat dipercaya untuk mengambil keputusan mengenai pencapaian kompetensi siswa.

Reabilitas

Reabilitas adalah suatu indikator yang menunjukkan seberapa konsisten sebuah pertanyaan. Dengan kata lain, reabilitas menunjukkan sejauh mana butir soal peranyaan dapat memberikan hasil yang serupa dalam pengujian yang dilakukan berulang kali (Rem et al., 2020). Perhitungan estimasi reabilitas dilakukan selama analisis butir soal, karena reabilitas juga berkaitan dengan suatu tes apakah sudah memenuhi kriteria yang ditetapkan. Hasil koefisien reabilitas dapat dilihat pada gambar 2.



Keterangan : Hasil Reabilitas Tes

Sumber : Data diolah 2025

Gambar 2. Hasil Reabilitas Tes.

Pada reabilitas data menunjukkan jumlah rata - rata nilai adalah 33,25 dan simpangan baku 14,36 dengan korelasi XY 0,95 dan reabilitas tes adalah 0,98. Tes tersebut dinilai handal apabila koefisien sekurang-kurangnya adalah 0,80. Berikut merupakan tabel kriteria korelasi koefisien :

Tabel 2. Kriteria Penafsiran Reabilitas Soal.

Nilai Reabilitas	Intepretasi
0,00 - 0,20	Korelasi sangat rendah
0,20 - 0,40	Korelasi rendah
0,40 - 0,70	Korelasi cukup
0,70 - 0,90	Korelasi tinggi
0,90 - 1,00	Korelasi sangat tinggi (Sempurna)

Sumber : Data diolah (2025).

Tingkat reliabilitas keseluruhan instrumen tes menunjukkan koefisien sebesar 0,97. Mengacu pada kriteria yang digunakan, koefisien ini termasuk dalam kategori "Korelasi sangat tinggi" atau mendekati "Sempurna" (rentang 0,90–1,00). Reliabilitas yang sangat tinggi ini menggarisbawahi bahwa instrumen tes memiliki tingkat kekonsistenan pengukuran yang sangat baik. Hal ini berarti tes tersebut mampu memberikan hasil yang stabil dan terpercaya jika diujikan kembali. Kekonsistenan hasil pengukuran merupakan syarat mutlak agar suatu instrumen evaluasi dapat dianggap handal (reliable) dalam konteks pengukuran hasil belajar.

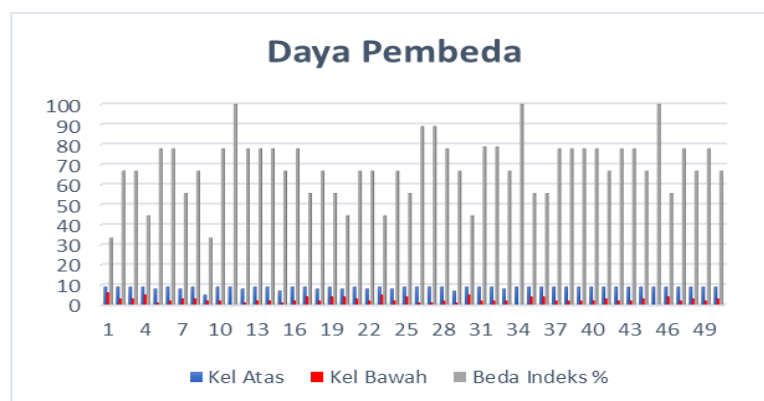
Daya Pembeda

Daya pembeda adalah seberapa efektif suatu soal dapat memisahkan peserta didik yang berada di kelompok atas dan bawah serta mereka yang memiliki kemampuan rendah. Ini dapat menggambarkan perbedaan yang signifikan dalam tingkat keberhasilan menjawab antara kedua kelompok tersebut (Hakim, 2025). Soal yang mampu membedakan siswa akan memberikan skor lebih tinggi kepada siswa yang memahami materi dengan baik, dan skor lebih rendah kepada siswa yang kurang memahami. Dibawah ini merupakan interpretasi dari koefisien daya pembeda :

Tabel 3. Interpretasi Koefisien Daya Pembeda.

Indeks Daya Pembeda	Interpretasi	Keterangan
0,00 - 0,20	Jelek	Soal diperbaiki
0,20 - 0,40	Cukup	Soal diterima dan diperbaiki
0,40 - 0,70	Baik	Soal diterima
0,70 - 1,00	Sangat baik	Soal diterima
Negatif	Semuanya tidak baik	Soal dibuang

Sumber : Data diolah (2025).



Keterangan : Hasil Daya Pembeda

Sumber : Data diolah 2025

Gambar 3. Hasil Daya Pembeda.

Distribusi butir soal berdasarkan daya pembeda ditunjukkan pada tabel 4.

Tabel 4. Hasil Distribusi Soal Daya Pembeda.

No.	Indeks Tingkat Kesukaran	No. Butir Soal	Jumlah	Presentasi
1	Jelek 0,00 - 0,20	-	-	-
2	Cukup 0,20 - 0,40	1, 9	2	4%
3	Baik 0,40 - 0,70	2, 3, 4, 7, 8, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25, 29, 30, 33, 35, 36, 41, 44, 46, 48, 50	25	50%
4	Sangat Baik 0,70 - 1,00	5, 6, 10, 11, 12, 13, 14, 16, 26, 27, 28, 31, 32, 34, 37, 38, 39, 40, 42, 43, 45, 47, 49	23	46%
5	Negatif	-	-	-
TOTAL			50	100 %

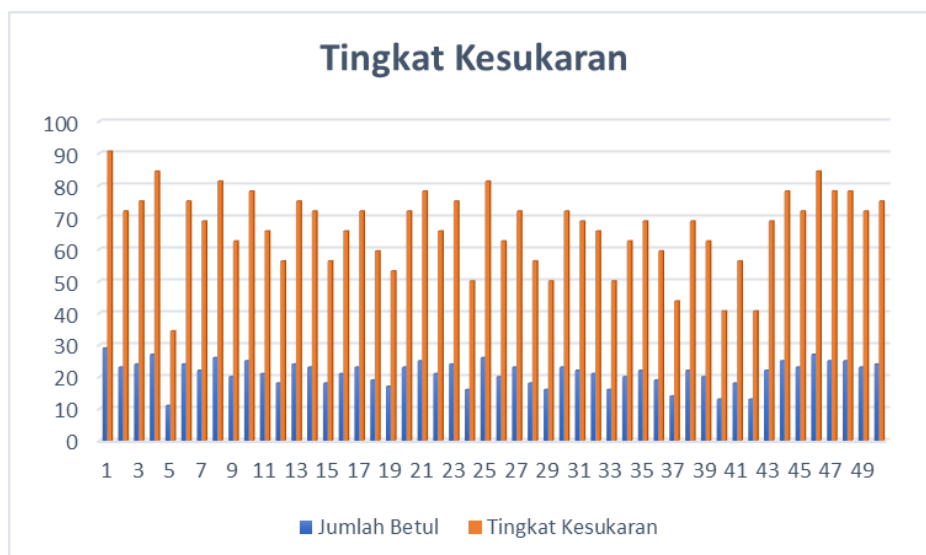
Sumber: Data diolah (2025).

Berdasarkan hasil analisis tabel 4, Hasilnya menunjukkan bahwa 96 % dari butir soal memiliki daya pembeda yang berfungsi baik atau sangat efektif. Diperoleh daya pembeda soal dengan kategori “Cukup” memiliki presentase 4% dengan total 2 soal, kategori “Baik” memiliki presentase 50% dengan total 25 soal, kategori “Sangat Baik” memiliki presentase 46% dengan total 23 soal.

Persentase yang sangat tinggi dari butir soal dengan daya pembeda yang baik dan sangat baik menunjukkan bahwa tes ini efektif dalam mengidentifikasi perbedaan kemampuan siswa. Hanya terdapat dua butir soal (4%), yaitu soal nomor 1 dan 9, yang tergolong kategori "Cukup". Butir-butir soal yang "Cukup" ini disarankan untuk dipertimbangkan perbaikannya agar kualitasnya optimal.

Tingkat Kesukaran

Dalam penilaian pendidikan, tingkat kesulitan adalah ukuran yang mencerminkan seberapa sulit atau seberapa mudah suatu soal bagi siswa. Indikator termasuk elemen penting untuk memastikan terdapat keseimbangan antara pertanyaan yang mudah, sedang, dan sulit. Tingkat kesukaran soal yang dianalisis menggunakan *software* Anates versi 4.0 ditunjukkan pada gambar 4.



Keterangan : Hasil Tingkat Kesukaran

Sumber : Data diolah 2025

Gambar 4. Hasil Tingkat Kesukaran.

Distribusi butir soal berdasarkan indeks kesukaran ditunjukkan pada tabel 5.

Tabel 5. Inteprestasi Tingkat.

No.	Indeks Tingkat Kesukaran	No. Butir Soal	Jumlah	Presentasi
1	Sangat Mudah 0,81 - 1,00	1	1	2 %
2	Mudah 0,71 - 0,80	2, 3, 4, 6, 8, 10 , 13, 14, 17, 20, 21, 23, 25, 27, 30, 44, 45, 46, 47, 48, 49, 50	22	44%
3	Sedang 0,31 - 0,70	5, 7, 9, 11, 12, 15, 16, 18, 19, 22, 24, 26, 28, 29, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43	27	54%
4	Sulit 0,16 - 0,30	-	-	-
5	Sangat Sulit 0,00 - 0,15	-	-	-
TOTAL			50	100 %

Sumber: Data diolah (2025).

Berdasarkan pada gambar 4 dan tabel 5 didapatkan indeks tingkat kesukaran yang menunjukkan bahwa dari keseluruhan soal dengan kategori “Sangat Mudah” memiliki presentase 2% dengan 1 soal, kategori “Mudah” memiliki presentase 44 % dengan jumlah 22 soal, dan kategori “Sulit” memiliki presentase 54 % dengan jumlah 27 soal. Analisis tingkat

kesukaran (difficulty level) menunjukkan komposisi butir soal sebagai berikut: Sedang 54% (27 soal), Mudah 44% (22 soal), dan Sangat Mudah 2% (1 soal).

Komposisi ini menunjukkan bahwa tes ini secara umum cenderung mudah bagi siswa SMK/MAK yang menjadi subjek penelitian, karena tidak ditemukan satu pun butir soal yang masuk kategori "Sulit" atau "Sangat Sulit". Idealnya, tes yang baik harus memiliki distribusi seimbang antara soal mudah, sedang, dan sulit. Ketiadaan soal sulit mengindikasikan bahwa tes ini belum maksimal dalam mengukur kemampuan berpikir tingkat tinggi (Higher Order Thinking Skills - HOTS) siswa. Implikasinya, perlu dilakukan pengembangan soal yang lebih menantang untuk mengukur pemahaman materi secara lebih komprehensif.

Kualitas Pengecoh (*Distractor*)

Distractor dalam soal pilihan ganda memiliki peran yang penting untuk mengkonfirmasi keabsahan dan konsistensi soal. *Distractor* yang baik tidak hanya membantu membedakan siswa yang memahami materi dengan yang tidak, tetapi juga meningkatkan akurasi dalam menilai pemahaman siswa terhadap materi yang diujikan (Hakim, 2025).

Kualitas pengecoh dalam analisis yang dilakukan dengan menggunakan *software* Anates menunjukkan bahwa setiap soal dilengkapi dengan pengecoh yang memiliki beragam tingkat kualitas (Puteri et al., 2025). Dalam Anates, total siswa yang memberikan respons untuk setiap pilihan dapat dengan mudah terlihat pada setiap opsi. Namun, identitas siswa yang memberi jawaban tersebut tidak dapat diidentifikasi. Berikut adalah tabel yang menunjukkan distribusi jawaban berdasarkan indeks pengecoh atau *distractor*, seperti yang terlihat pada tabel 6.

Tabel 6. Distribusi Jawaban Indeks Pengecoh.

No	Pengecoh	Jumlah	Presentasi
1	Sangat Baik (+ +)	61	30,5%
2	Baik (+)	45	22,5%
3	Kurang Baik (-)	44	22%
4	Buruk (- -)	38	19%
5	Sangat Buruk (- - -)	12	6%
TOTAL		200	100 %

Sumber: Data diolah (2025).

Berdasarkan pada tabel 6 didapatkan distribusi soal pada indeks pengecoh menunjukkan bahwa dari keseluruhan butir soal jawaban pengecoh dengan kategori “Sangat Baik” memiliki presentase 30,5% sejumlah 61 jawaban, kategori “Baik” memiliki presentase 22,5 % sejumlah 45 jawaban, kategori “Kurang Baik” memiliki presentase 22 % sejumlah 44

jawaban, kategori “Buruk” memiliki presentase 19 % sejumlah 38 jawaban, dan kategori “Sangat Buruk” memiliki presentase 6 % sejumlah 12 jawaban.

Dari hasil tersebut, jawaban dengan kategori “Buruk” dan “Sangat Buruk” perlu diperbaiki agar lebih efektif. Sementara jawaban dengan kategori “Kurang” “Baik” dan “Sangat Baik” sudah dapat dikatakan cukup dan memenuhi kriteria. Pengecoh yang baik dapat membantu pendidik dalam mengetahui tingkat pemahaman siswanya, bukan karena proses eliminasi yang mudah.

Kualitas pengecoh (distractor) dinilai berdasarkan efektivitasnya dalam menarik perhatian siswa yang tidak mengetahui jawaban benar. Hasil analisis (Tabel 6) menunjukkan bahwa 53% dari total pengecoh berfungsi "Sangat Baik" (++) atau "Baik" (+). Namun, hampir separuhnya (47%) masih tergolong "Kurang Baik" (-) "Buruk" (--), atau "Sangat Buruk" (---).

Pengecoh yang berfungsi buruk (termasuk yang "Buruk" dan "Sangat Buruk") adalah opsi jawaban yang tidak dipilih sama sekali atau dipilih sangat sedikit oleh kelompok siswa yang kemampuannya rendah. Kondisi ini membuat soal menjadi terlalu mudah karena siswa dapat dengan mudah mengeliminasi pilihan jawaban yang salah, bukan karena penguasaan materi yang mendalam. Oleh karena itu, pengecoh yang kurang efektif perlu direvisi agar daya tarik opsinya meningkat, sehingga meningkatkan kualitas diskriminasi soal.

5. KESIMPULAN DAN SARAN

Penelitian ini mengungkapkan bahwa mutu butir soal dalam Ekonomi Bisnis dan Administrasi Umum untuk siswa kelas XI dan XII di SMK Akuntansi sudah memenuhi standar instrumen evaluasi yang baik berdasarkan analisis Anates V4. Semua soal dinyatakan valid, instrumen menunjukkan tingkat reliabilitas yang sangat tinggi, dan sebagian besar butir soal memiliki kemampuan membedakan serta tingkat kesulitan yang tepat. Pengecoh berfungsi dengan cukup baik meski ada beberapa yang masih memerlukan perbaikan. Hasil ini menunjukkan bahwa instrumen penilain mendukung asesmen yang tepat dalam kerangka Kurikulum Merdeka.

Namun, penelitian ini belum melakukan analisis mendalam tentang kesesuaian soal dengan indikator pencapaian pembelajaran maupun metode penyusunan soal oleh guru, sehingga aspek pedagogis masih menjadi area yang belum dibahas. Untuk penelitian berikutnya, disarankan agar dilakukan evaluasi terhadap penerapan hasil analisis instrumen dalam pengambilan keputusan pembelajaran, serta mengembangkan model pendampingan bagi guru dalam menyusun soal menggunakan perangkat lunak. Penelitian lanjutan juga bisa

mencoba pendekatan analisis alternatif untuk membuat hasil evaluasi instrumen lebih komprehensif.

DAFTAR REFERENSI

- Adimas, I., Saputra, G., Hakim, L., Pratiwi, V., Rohayati, S., Wahjudi, E., & Farid, M. M. (2025). Pelatihan analisis butir soal berbasis ANATES untuk meningkatkan kualitas evaluasi pembelajaran guru MGMP Akuntansi. *Jurnal Pengabdian kepada Masyarakat*, 6(3), 1288-1297.
- Agustini, N., Mutiara, A., Yulaini, E., & Aradea, R. (2021). Analisis kemampuan siswa dalam menyelesaikan soal-soal persediaan metode FIFO, LIFO, dan moving average. *Journal of Education Research*, 2(2), 69-73. <https://doi.org/10.37985/jer.v2i2.49>
- Anazar, & Munir, M. (2024). Analisis kualitas butir soal ujian akhir semester genap mata pelajaran simulasi dan komunikasi digital. *Journal of Information Technology and Education*, 2(1), 41-51. <https://doi.org/10.21831/jited.v2i1.275>
- Andriatoni, Adia, B., Yessari, M., & Yeni, Z. (2025). Asesmen otentik sebagai sarana penguatan Profil Pelajar Pancasila. *Jurnal Pendidikan*, 11, 1-8.
- Anike, A. (2022). Analisis kualitas butir soal ujian tengah semester (UTS) ganjil biologi kelas XI SMAN 1 Pancung Soal. *Jurnal Pendidikan Biologi*, 4(2), 61-69. <https://doi.org/10.37058/metaedukasi.v4i2.4087>
- Artama, S., Djollong, I., Lubis, K., Yulianti, M., Mardin, I., Fatih, H., & Holifah, D. (2023). Evaluasi hasil belajar. PT Mifandi Mandiri Digital.
- Asyarah, S. P., Febriyanti, A. C., & Sitanggang, C. J. C. (2025). Analisis butir soal HOTS pada elemen perpajakan fase F SMK Akuntansi menggunakan software ANATES. *Jurnal Pendidikan Akuntansi*, 4(2), 1594-1602.
- Augustia, A. D., & Agustia, C. N. (2025). Analisis validitas dan reliabilitas soal pilihan ganda menggunakan software ANATES pada mata pelajaran perpajakan. *Jurnal Pendidikan Akuntansi*, 5(1), 250-265. <https://doi.org/10.37481/jmneb.v5i1.1165>
- Damayanti, F. A., Fauziah, A. N., Hakim, L., & Pratiwi, V. (2024). Analisis butir soal HOTS pilihan ganda siswa kelas X SMK Negeri 1 Jombang menggunakan aplikasi ANATES. *Jurnal Pendidikan*, 8, 48686-48693.
- Fietri, W. A., Lufri, Syamzurizal, & Zulyusri. (2021). Analisis butir soal biologi kelas VIII Madrasah Tsanawiyah Negeri 6 Kerinci. *Jurnal Pendidikan Biologi Undiksha*, 8, 1-10.
- Firandhika, A., Apriliyanti, K. D., Kasih, A., & Saputri, A. (2024). Analisis butir soal HOTS pilihan ganda pada elemen komputer akuntansi di SMKN 10 Surabaya menggunakan aplikasi ANATES. *Jurnal Pendidikan Akuntansi*, 4(1), 727-738.
- Fiska, J. M., Hidayati, Y., Qomaria, N., & Hadi, W. P. (2021). Analisis butir soal ulangan harian IPA menggunakan software ANATES berdasarkan pendekatan teori tes klasik. *Jurnal Pendidikan IPA*, 4(1), 1-10. <https://doi.org/10.21107/nser.v4i1.8133>
- Fitraynsyah, A., & Hilmiyati, F. (2024). Peran tingkat kesukaran dan daya pembeda dalam analisis butir tes: Kajian literatur untuk pendidikan menengah. *Jurnal Evaluasi Pendidikan*, 1(4), 252-262.
- Hakim, L. (2025). Analisis butir soal elemen ekonomi bisnis dan administrasi umum

- menggunakan aplikasi ANATES. *Jurnal Pendidikan Akuntansi*, 5(1), 1-12.
- Ikawati, D. S., Agustin, E. W., Hakim, L., & Pratiwi, V. (2024). Analisis butir soal pilihan ganda siswa SMK menggunakan ANATES. *Edukatif: Jurnal Ilmu Pendidikan*, 6(6), 6773-6781.
- Iriyanti, R. D., Agustina, N. D., & Prinita, D. A. (2024). Analisis butir soal pilihan ganda mata pelajaran ekonomi bisnis dan administrasi umum menggunakan software ANATES di SMK NU Gresik. *Jurnal Pendidikan Akuntansi*, 4(2), 595-613. <https://doi.org/10.56689/padma.v4i2.1573>
- Jatmoko, D., Pambudi, A., Setiyono, K. A., & Nugroho, B. (2025). Analisis butir soal teknik dasar otomotif berbasis diagnostik kognitif. *Jurnal Pendidikan Teknik Mesin*, 4(1), 1-10.
- Karim, N. (2024). Pengaruh teman sebaya dan pemanfaatan fasilitas belajar terhadap hasil belajar ekonomi bisnis dan administrasi umum. *Indonesian Research Journal on Education*, 4, 2510-2523.
- Manik, P., Saraswati, S., & Agustika, G. N. (2020). Kemampuan berpikir tingkat tinggi dalam menyelesaikan soal HOTS mata pelajaran matematika. *Jurnal Pendidikan Matematika*, 4(2), 257-269. <https://doi.org/10.23887/jisd.v4i2.25336>
- Prinita, D. A., Iriyanti, R. D., Agustina, N. D., & Arif, A. (2024). Pengembangan handout digital berbasis flipbook pada materi ekonomi bisnis dan administrasi umum. *Jurnal Pendidikan Akuntansi*, 4(2), 614-625. <https://doi.org/10.56689/padma.v4i2.1572>
- Purwati, L. M., Arianty, R., Syakilah, D. M., & Ridlo, S. (2021). Analisis soal tes pilihan ganda berbasis HOTS menggunakan aplikasi ANATES. *Jurnal Evaluasi Pendidikan*, 15, 460-473. <https://doi.org/10.52434/jp.v15i2.1287>
- Puteri, N. D., Arsyanti, K., Putricia, N. D., Hakim, L., & Pratiwi, V. (2025). Analisis kelayakan butir soal pada handout akuntansi keuangan dana kas kecil berbasis e-book menggunakan ANATES. *Jurnal Pendidikan Akuntansi*, 2(3), 205-214. <https://doi.org/10.62421/jibema.v2i3.108>
- Rachmadani, F. C., Hakim, L., & Pratiwi, V. (2024). Analisis butir soal pilihan ganda pada elemen akuntansi keuangan guna mengoptimalkan evaluasi. *Jurnal Nirta: Studi Inovasi*, 4, 1-18.
- Rem, S. (2020). Analisis kualitas butir soal tipe HOTS pada kompetensi sistem rem siswa SMK. *Jurnal Pendidikan Vokasional*, 17(1), 95-104. <https://doi.org/10.23887/jptk-undiksha.v17i1.22914>
- Syafiriya, D. M., & Hakim, L. (2024). Analisis butir soal HOTS untuk menguji pemahaman siswa SMK jurusan akuntansi menggunakan software ANATES. *Jurnal Pendidikan Akuntansi*, 4(1), 1-10.