

## ANALISIS BUTIR SOAL PILIHAN GANDA PADA ELEMEN AKUNTANSI KEUANGAN GUNA MENGOPTIMASI EVALUASI MENGUNAKAN ANATES V4.0

**Febilia Ceril Rachmadani**  
Universitas Negeri Surabaya

**Cahyani Dwi Lestari**  
Universitas Negeri Surabaya

**Luqman Hakim**  
Universitas Negeri Surabaya

**Vivi Pratiwi**  
Universitas Negeri Surabaya

Korespondensi penulis: [febilia.23012@mhs.unesa.ac.id](mailto:febilia.23012@mhs.unesa.ac.id), [cahyani.23091@mhs.unesa.ac.id](mailto:cahyani.23091@mhs.unesa.ac.id),  
[luqmanhakim@unesa.ac.id](mailto:luqmanhakim@unesa.ac.id), [vivipratiwi@unesa.ac.id](mailto:vivipratiwi@unesa.ac.id)

**Abstract.** *Improving the quality of assessment instruments is a critical aspect of the modern education system, especially in the context of vocational education in Indonesia. This study focuses on the optimization of multiple-choice items on the financial accounting element in Vocational High Schools (SMK) using Anates V4 software, a sophisticated test analysis instrument that allows in-depth evaluation of the quality of assessment instruments. Using a quantitative descriptive approach, the study involved 20 students of class XI majoring in accounting and analyzed 10 multiple-choice items through a series of quality parameters, including reliability, validity, distinguishing power, and difficulty level. The analysis was conducted by utilizing the comprehensive capabilities of Anates V4 to provide a comprehensive assessment of the evaluation instrument. The results showed that the reliability coefficient of the test was 0.77, including the high reliability category. In terms of validity, 6 out of 10 questions had a significant correlation, while 4 questions required revision or deletion. Differentiating power analysis revealed that 8 questions were categorized as good to very good, while 2 questions needed improvement. The difficulty level of the questions varied, with 1 easy question, 8 moderate questions, and 1 very difficult question. This study confirms that the use of Anates V4 provides a comprehensive analysis of the quality of test instruments, assisting educators in identifying and improving items to increase the effectiveness of assessment in learning financial accounting.*

**Keywords:** *Anates V4.0, Evaluation, Item Analysis, Item Optimization*

**Abstrak.** Peningkatan kualitas instrumen evaluasi merupakan aspek kritis dalam sistem pendidikan modern, khususnya dalam konteks pendidikan kejuruan di Indonesia. Penelitian ini fokus pada optimasi butir soal pilihan ganda pada elemen akuntansi keuangan di Sekolah Menengah Kejuruan (SMK) dengan menggunakan software Anates V4, sebuah instrumen analisis tes canggih yang memungkinkan evaluasi mendalam terhadap kualitas instrumen penilaian. Menggunakan pendekatan deskriptif kuantitatif, penelitian melibatkan 20 siswa kelas XI jurusan akuntansi dan menganalisis 10 butir soal pilihan ganda melalui serangkaian parameter kualitas, mencakup reliabilitas, validitas, daya pembeda, dan tingkat kesukaran. Analisis dilakukan dengan memanfaatkan kemampuan komprehensif Anates V4 untuk memberikan penilaian menyeluruh terhadap instrumen evaluasi. Hasil penelitian menunjukkan bahwa koefisien reliabilitas tes sebesar 0,77, termasuk kategori reliabilitas tinggi. Dari segi validitas, 6 dari 10 soal memiliki korelasi signifikan, sementara 4 soal memerlukan revisi atau penghapusan. Analisis daya pembeda mengungkapkan 8 soal berkategori baik hingga sangat baik, sedangkan 2 soal membutuhkan perbaikan. Tingkat kesukaran soal bervariasi, dengan 1 soal mudah, 8 soal berkategori sedang, dan 1 soal

sangat sukar. Penelitian ini menegaskan bahwa penggunaan Anates V4 memberikan analisis komprehensif terhadap kualitas instrumen tes, membantu pendidik dalam mengidentifikasi dan memperbaiki butir soal guna meningkatkan efektivitas penilaian dalam pembelajaran akuntansi keuangan.

**Kata kunci:** Analisis Butir Soal, Anates V4.0, Evaluasi, Optimasi Butir Soal

## **LATAR BELAKANG**

Sektor pendidikan menjadi pilar utama dalam upaya pengembangan dan kemajuan suatu bangsa, dengan kontribusi yang sangat signifikan dalam pembangunan berkelanjutan. (Drs. Syafril, 2019) menegaskan bahwa pendidikan memiliki fungsi strategis dalam mempersiapkan peserta didik menghadapi dinamika perubahan lingkungan yang sangat cepat, serta memberi motivasi untuk berperan aktif dalam transformasi kehidupan menuju arah yang lebih positif.

Di Sekolah Menengah Kejuruan, peserta didik mendapatkan pengetahuan dan keterampilan melalui dua jenis pembelajaran yakni teori dikelas dan praktik di ruang praktikum (Warju, Ariyanto, Soeryanto, & Trisna, 2020). Setelah proses pembelajaran selesai, para pendidik melakukan penilaian guna mengkaji kemampuan peserta didik untuk mendapatkan penilaian yang akurat dan diperlukan instrumen yang valid serta objektif. Evaluasi merupakan sebuah rangkaian kegiatan yang tak terhindarkan dan sangat diperlukan dalam menunjang berjalannya proses pendidikan secara menyeluruh (Prastika, 2021). Menurut (Riani, Almujab, Dina, & Budiarto, 2020) untuk mengoptimalkan mutu pendidikan, diperlukan keberadaan komponen-komponen pendukung yang integral, di mana evaluasi berperan sebagai mekanisme kunci dalam mendorong pengembangan sistem pendidikan secara berkelanjutan. Aktivitas evaluasi meliputi proses sistematis yang terdiri dari pengukuran yakni membandingkan objek dengan instrumen yang telah dirancang serta penilaian yang didasarkan pada interpretasi komprehensif dari hasil pengukuran tersebut (Elviana, 2020a). Dengan proses evaluasi, mutu pembelajaran yang sudah dilakukan oleh pendidik dapat berdampak pada peningkatan kemampuan peserta didik. Peningkatan kapasitas dan kompetensi siswa memiliki korelasi langsung dengan mutu pendidikan, di mana semakin tinggi kemampuan peserta didik, maka semakin berkualitas pula sistem pendidikan yang bersangkutan (Khasanah, Fuady, & Sunismi, 2023).

Seiring dalam perkembangan di dunia pendidikan, terutama pada proses evaluasi pembelajaran yang memiliki peran kritis, terutama dalam menilai efektifitas pengajaran

dan tingkat pemahaman peserta didik. Upaya peningkatan mutu pendidikan di Indonesia perlu di dilaksanakan agar sistem pendidikan dapat responsif dan fleksibel menghadapi perkembangan global. Guna mewujudkan pendidikan berkualitas, diperlukan berbagai komponen pendukung, di mana asesmen belajar berperan dalam mendorong kemajuan dan transformasi dunia pendidikan.

Asesmen belajar pada dasarnya adalah cara untuk mengumpulkan berbagai informasi yang diperlukan guna membuat keputusan yang tepat berkaitan dengan proses dan kemajuan peserta didik (Azzahroh, Iman, Anwar, & Aziz, 2022). Dalam sebuah proses evaluasi, tes menjadi alat pengukur yang pakai untuk menimbang hasil belajar peserta didik. Tes sendiri dapat didefinisikan sebagai sebuah metode atau langkah-langkah sistematis yang dijalankan untuk melakukan pengukuran dan penilaian dalam konteks pendidikan. Selain itu (Ismanto, 2014) (Mania, Fitriani, Majid, Ichiana, & Abrar, 2020) menyatakan bahwa tes merupakan instrumen evaluasi dalam pendidikan yang terdiri dari tugas berupa pertanyaan atau perintah yang harus diselesaikan oleh siswa sebagai alat ukur pencapaian pembelajaran. Menurut (Sudijono, 2009)(Muluki, 2020), tes mempunyai dua fungsi utama: 1) Mengukur kompetensi peserta didik yang diperoleh selama proses pembelajaran, dan 2) Mengevaluasi keberhasilan program pengajaran dengan mengukur efektivitas pelaksanaan pembelajaran yang telah dilakukan. Diantara berbagai jenis tes yang ada, tes pilihan ganda adalah salah satu bentuk asesmen yang populer dan memerlukan analisis mendalam untuk memastikan kualitas dan efektifitasnya dalam mengukur pemahaman peserta didik terhadap elemen-elemen akuntansi keuangan. Melalui proses analisis yang mendalam, tes ini berpotensi menjadi instrumen evaluasi yang mampu menghasilkan penilaian objektif dan presisi. Namun, apabila tes tersebut diimplementasikan oleh tenaga pendidik, maka hasil yang diperoleh diperkirakan tidak akan memenuhi kualitas yang diharapkan (Fietri, Lufri, Syamzurizal, & Zulyusri, 2021)

Seiring dengan implementasi Kurikulum Merdeka yang mengedepankan pendekatan fleksibel dan berbasis kompetensi, peningkatan mutu instrumen evaluasi menjadi semakin krusial dalam lanskap pendidikan Indonesia kontemporer. Sistem evaluasi tidak hanya dituntut untuk selaras dengan standar kurikulum, tetapi juga harus dapat mengakomodasi keragaman kapasitas dan kebutuhan peserta didik di berbagai wilayah. Kemajuan teknologi pendidikan membuka kesempatan untuk mengadopsi

sistem analisis digital dalam pengembangan dan pengelolaan instrumen evaluasi. *Software* ANATES hadir sebagai instrumen yang dapat diandalkan untuk mengoptimalkan mutu butir soal melalui serangkaian analisis yang terstruktur dan menyeluruh. ANATES versi 4.0 for windows dirancang secara khusus guna melakukan analisis terhadap soal-soal pilihan ganda dan esai (Maulidah, Sukarno, & Syefrinando, 2023). *Software* ini merupakan hasil pengembangan Karnoto dan Yudi Wibisono, yang resmi dirilis pada tanggal 22 Februari 2004 silam.

Analisis butir soal merupakan sebuah metode terstruktur yang bertujuan mengungkapkan informasi terperinci dan spesifik mengenai komponen-komponen dalam suatu instrumen tes yang telah disusun (Surapranata, Sumarna, 2006). Maksud dilakukannya analisis butir soal adalah untuk mendapatkan pemahaman mendalam tentang ciri-ciri spesifik dari setiap item pertanyaan, dengan menggunakan pendekatan kajian teoritis dan pengujian empiris (Gusmizain, 2022). Pengaplikasian ANATES dalam analisis butir soal memberikan kesempatan bagi para tenaga pendidik untuk melakukan evaluasi mendalam terhadap beragam komponen esensial seperti validitas, reliabilitas, indeks kesukaran, kemampuan pembeda, serta fungsi pengecoh pada tiap butir soal. Menurut (Lasmy, Yusrizal, & Razali, 2021) Menganalisis kualitas butir soal adalah langkah penting yang wajib ditempuh oleh para pendidik guna mengoptimalkan dan menyempurnakan mutu instrumen penilaian yang telah disusun. Urgensi optimasi ini terletak pada kebutuhan untuk mengembangkan perangkat evaluasi yang dapat memberikan pemetaan akurat tentang pencapaian kompetensi peserta didik dalam ranah akuntansi keuangan. Melalui proses analisis dan optimasi yang terstruktur, diharapkan dapat terwujud instrumen evaluasi yang lebih tepat guna dalam mengukur hasil pembelajaran. Proses optimasi menggunakan ANATES tidak terbatas pada pengolahan data kuantitatif semata, namun juga meliputi analisis efektivitas pengecoh yang memberikan masukan berharga mengenai kualitas opsi jawaban yang tersedia. Pendekatan yang holistik ini memfasilitasi pendidik dalam melakukan penyempurnaan yang terarah, sehingga menghasilkan instrumen evaluasi yang lebih terpercaya dalam mengukur capaian kompetensi peserta didik di bidang akuntansi keuangan. Dengan demikian, implementasi ANATES dalam pengoptimalan butir soal pilihan ganda terbukti menjadi pendekatan yang tepat untuk meningkatkan akurasi pengukuran kompetensi dalam pembelajaran akuntansi keuangan.

## **KAJIAN TEORITIS**

Evaluasi pendidikan merupakan konstruk fundamental dalam arsitektur sistemik pendidikan yang memiliki fungsi strategis untuk mengidentifikasi efektivitas proses pembelajaran dan mengukur capaian kompetensi peserta didik. Sebagai instrumen diagnostik yang komprehensif, evaluasi tidak sekadar mekanisme pengukuran konvensional, melainkan pendekatan sistematis yang mempertimbangkan parameter teknis seperti validitas, reliabilitas, indeks kesukaran, kemampuan pembeda, serta fungsi pengecoh pada tiap butir soal. Instrumen evaluasi, khususnya tes pilihan ganda, mensyaratkan analisis multidimensional untuk menjamin akurasi dan representativitas hasil pengukuran pendidikan.

Perkembangan teknologi informasi telah mentransformasi metodologi analisis instrumen evaluasi. Perangkat lunak seperti Anates V4 mereresentasikan inovasi teknologis dalam domain evaluasi pendidikan, yang memungkinkan pengkajian kuantitatif dan kualitatif terhadap instrumen pengukuran secara komprehensif dan presisi. Validitas dan reliabilitas merupakan konstruk metodologis fundamental dalam epistemologi pengukuran pendidikan, yang secara komprehensif menentukan kualitas instrumen evaluasi. Validitas didefinisikan sebagai kapasitas instrumen untuk mengukur konstruk yang dimaksudkan dengan akurasi, sementara reliabilitas mengindikasikan konsistensi dan reproduksibilitas hasil pengukuran dalam berbagai kondisi empiris. Analisis daya pembeda dan tingkat kesukaran soal merupakan prosedur metodologis kritis dalam mengonstruksi instrumen evaluasi berkualitas. Pendekatan ini memungkinkan diferensiasi komprehensif kemampuan peserta didik melalui identifikasi karakteristik psikometrik setiap butir soal, yang pada gilirannya memberikan kontribusi signifikan terhadap keakuratan penilaian kompetensi akademik.

## **METODE PENELITIAN**

Penelitian analisis butir soal ini termasuk dalam penelitian evaluatif dengan pendekatan deskriptif kuantitatif. Artinya, penelitian kuantitatif ini dilaksanakan bukan untuk membuktikan atau menyangkal hipotesis. Namun, untuk mengungkapkan dan menggambarkan kondisi objektif sesuai dengan fenomena yang sedang dikaji secara transparan dan faktual. Hasil dari responden yang berupa angka akan dianalisis dengan menggunakan *software* Anates V4. Sesudah dianalisis kemudian diinterpretasikan secara deskriptif berupa persentase. Penelitian bertujuan menilai kualitas item berbasis HOTS

yang dibuat dan diberikan kepada siswa SMK kelas xi dengan menilai validitas, reliabilitas, tingkat kesulitan, dan daya beda dengan melalui kuisisioner berisikan 10 soal pilihan ganda (*multiple choice*).

Subjek penelitian merupakan siswa Sekolah Menengah Kejuruan kelas XI jurusan akuntansi. Subjek penelitian diambil secara acak (*random sampling*). Kemudian didapatkan responden sebanyak 20 siswa yang secara sukarela mengisi kuesioner tersebut. Penelitian ini menerapkan instrumen evaluasi berbentuk tes objektif dengan format pilihan ganda (*multiple choice*) dengan 10 soal, masing-masing memiliki 5 opsi jawaban dari A - E. Tes objektif merupakan metode evaluasi yang menggunakan format pilihan ganda, di mana setiap pertanyaan dilengkapi dengan sejumlah alternatif jawaban yang memungkinkan peserta memilih respons yang dianggap benar (Fietri, 2021). Pengumpulan data dalam penelitian ini dilakukan melalui instrumen berbentuk kuesioner elektronik (google form), kemudian dibagikan kepada peserta didik kelas XI jurusan akuntansi. Instrumen tes tersebut mencakup materi spesifik tentang konsep utang dan piutang dalam elemen akuntansi keuangan kelas xi. Dari hasil responden tersebut akan dievaluasi melalui aplikasi *software* Anates V4.0. (Fiska, Hidayati, Qomaria, & Hadi, 2021) menekankan bahwa penilaian kualitas butir soal memerlukan analisis menyeluruh terhadap empat kriteria utama yang mencakup reliabilitas butir soal, validitas butir soal, kemampuan pembeda butir soal, dan tingkat kesulitan atau kesukaran butir soal. Hasil pengujian soal akan dipaparkan secara narasi deskriptif, dengan mengintegrasikan data numerik yang diperoleh melalui perhiyungan menggunakan aplikasi *software* Anates V4.0.

## HASIL DAN PEMBAHASAN

Optimasi butir soal pada tes objektif berbentuk pilihan ganda (*multiple choice*) pada elemen akuntansi keuangan ini menggunakan *software* Anates, supaya mendapatkan hasil yang lebih akurat dalam menganalisis butir soal yang akan diujikan. Anates (Analisis Tes) merupakan sebuah program aplikasi pada komputer yang dimana khusus dikembangkan sebagai program yang dapat menganalisis kualitas butir soal dan tes secara keseluruhan. Dengan program ini sangat membantu para pendidik dan evaluator saat melakukan analisis pada instrumen penilaian dengan lebih efisien dan akurat (Eprillison & Gumanti, n.d). Anates memiliki kemampuan untuk menganalisis beberapa aspek penting dalam evaluasi butir soal. Pertama, program Anates ini menganalisis validitas

item soal melalui perhitungan hubungan antara skor individual dan skor keseluruhan. Dalam menganalisis reliabilitas, Anates menggunakan formula statistik untuk menghitung konsistensi internal tes. Anates juga menganalisis daya pembeda soal dengan membandingkan jawaban kelompok berkemampuan tinggi dan rendah. Pada tingkat kesukaran yang dianalisis dengan Anates secara teknis diukur melalui rasio keberhasilan peserta dalam menjawab soal dibandingkan keseluruhan peserta. Pada penelitian ini analisis dari butir soal akuntansi keuangan pada kelas xi yang telah diinput pada *software* Anates V4 menghasilkan analisis, yakni:

### **Reliabilitas**

Keandalan atau tingkat kepercayaan menilai kemampuan instrumen dalam memberikan hasil yang dapat diandalkan merupakan definisi dari Kamus Besar Bahasa Indonesia atau KBBI. Reliabilitas dalam konteks pengukuran, termasuk tes dapat diartikan pada menentukan kapasitas instrumen untuk menghasilkan pengukuran yang stabil dan terpercaya (Magdalena, Fauziah, Faziah, & Nopus, 2021). Menurut (Fakhri Ramadhan, Siroj, & Win Afgani, 2024) Reliabilitas menggambarkan kemampuan suatu alat ukur untuk memberikan hasil penilaian yang ajeg dan tepat, dimana penggunaan berulang akan menghasilkan skor yang konsisten dan dapat dipercaya.

Menurut (Akhmadi, 2021) karakteristik utama suatu tes yang reliabel tercermin melalui empat aspek, yaitu akurasi pengukuran, ketepatan prediksi, konsistensi hasil, dan stabilitas instrumen. Semakin baik keempat aspek tersebut, semakin tinggi pula reliabilitas yang digunakan. Selanjutnya, tingkat reliabilitas dapat diidentifikasi melalui koefisien korelasi antara skor yang diperoleh dari instrumen yang memiliki kesetaraan. Instrumen dikatakan reliabel manakala mampu memberikan hasil pengukuran yang secara substansial mendekati kondisi sesungguhnya dari objek yang diukur (Mardiyah & Rofi, 2024)

Kriteria reliabilitas merupakan aspek fundamental dalam mengukur kualitas instrumen penelitian yang mencerminkan tingkat konsistensi dan keajegan suatu alat ukur. Kriteria untuk mengukur reliabilitas sebagai berikut :

**Tabel.1 Kriteria Reliabilitas**

Kriteria Nilai	Kategori
$0,00 \leq r_{11} \leq 0,20$	Sangat rendah
$0,20 \leq DB \leq 0,40$	Rendah
$0,40 \leq DB \leq 0,70$	Cukup
$0,70 \leq DB \leq 0,90$	Tinggi
$0,90 \leq DB \leq 1,00$	Sangat tinggi

Pada penelitian ini butir soal elemen akuntansi keuangan kelas xi SMK yang telah dianalisis menggunakan software Anates V4, menunjukkan bahwa koefisien korelasi yaitu 0,62, sedangkan untuk koefisien reliabilitas tes mendapatkan sebesar 0,77. Tingkat kehandalan tes ditentukan dengan koefisien korelasi minimal 0,80. Maka dari itu, dapat diartikan bahwa nilai koefisien korelasi belum dapat sepenuhnya dianggap handal dan membutuhkan evaluasi lebih mendalam. Selanjutnya, nilai koefisien reliabilitas tes menurut tabel diatas kriteria reliabilitas tes dapat diandalkan karena memiliki kemandapan hasil yang signifikan.

No. Urut	No. Subyek	Kode/Nama Subyek	Skor Sampel	Skor Kelas	Skor Total
1	1	Givi Karyah	3	0	3
2	2	Rizka Mulya	2	1	3
3	3	Rizka Ananda Putri	1	1	2
4	4	Alfara Zakaria Ghita Agum	5	2	5
5	5	Ahmad Widiyadharma P	3	1	4
6	6	Suzana R	2	2	4
7	7	Alya Nurul Syafiq	0	0	0
8	8	Della Tereza Putri	3	2	5
9	9	Arissa Devi Hassan	5	1	6
10	10	Yusron Ad Kusnawan	1	0	1
11	11	Indah Widiyati	2	4	6
12	12	Azzam N	1	2	3
13	13	M. Adnan Pramana Putra	3	0	3
14	14	Shafiq Hidayat	5	4	9
15	15	Rara Putri Nisa	5	3	8
16	16	Anggun Puspa	5	3	8
17	17	Diki Anggi Rafi	4	2	6
18	18	Dewi Anugrah R	5	5	10
19	19	Andri Akwaningsih	5	4	9
20	20	Ayu Ningsih Wulandari	5	4	9

**Gambar.1 Hasil Reliabilitas**

## Validitas Butir Soal

Validitas tes adalah suatu konsep yang menggambarkan kemampuan suatu instrumen atau alat ukur untuk mengevaluasi objek pengukuran dengan akurasi maksimal. Dalam konteks evaluasi dan penelitian, validitas merupakan salah satu kriteria terpenting untuk menentukan kualitas dan ketepatan suatu instrumen pengukur (Prastika, 2021) Validitas pada dasarnya yaitu keabsahan atau ketepatan suatu instrumen dalam proses menilai target pengukuran yang spesifik. Terdapat tiga metode validasi standar: validitas isi, konstruk, dan empiris. Khusus pada validitas isi, kedalaman pengukuran ditentukan oleh relevansi dan keterwakilan setiap butir soal terhadap karakteristik atau konsep yang menjadi fokus penelitian. Validitas konstruk menyoroti aspek psikologis yang mencakup spektrum kemampuan kognitif, dimulai dari penguasaan informasi sederhana hingga keterampilan berpikir tingkat lanjut yang melibatkan analisis dan sintes (Syamsurizal, n.d.) .Validitas empiris diperoleh melalui pengujian instrumen pada sekelompok responden yang memiliki karakteristik serupa dengan populasi target guna memastikan ketepatan dan representativitas hasil pengukuran(Fakhri Ramadhan et al., 2024).

Menurut (Mardiyah & Rufi, 2024) validitas item mengukur keterkaitan hubungan antara pencapaian tiap item dengan performa menyeluruh. Proses ini dilakukan melalui perhitungan koefisien korelasi yang menggambarkan seberapa signifikan kontribusi item tersebut dalam mengukur konstruk atau variabel yang dimaksud. Semakin tinggi koefisien korelasi, semakin valid item tes dalam mewakili keseluruhan instrumen pengukuran. Tingkat signifikansi yang biasanya digunakan adalah 0,05 dan 0,01. Tingkatan ini dapat menentukan layak tidaknya item tes tersebut. Validitas suatu item dapat dibuktikan jika terdapat korelasi signifikan antara item tersebut dengan skor total. Batas signifikansi koefisien korelasi dapat dilihat sebagai berikut:

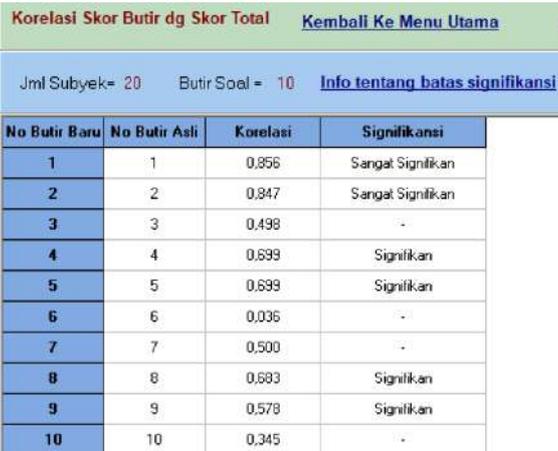
Catatan: Batas signifikansi koefisien korelasi sebagai berikut:

df (N-2)	P=0,05	P=0,01	df (N-2)	P=0,05	P=0,01
10	0,576	0,708	60	0,250	0,325
15	0,482	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	0,273	0,354	>150	0,159	0,208

Bila koefisien = 0,000 berarti tidak dapat dihitung.

**Gambar.2 Batas signifikansi koefisien korelasi**

Berdasarkan analisis yang telah dilakukan menggunakan *software* anates V4 pada butir soal pilihan ganda elemen akuntansi keuangan kelas xi SMK menunjukkan bahwa sebagian besar item soal memperlihatkan hubungan statistik yang kuat dengan nilai keseluruhan. Artinya, pada sebagian besar butir soal berkontribusi secara signifikan dalam mengukur konstruk yang sama dengan seluruh tes. Selain itu juga ada beberapa item soal tanpa korelasi signifikan. Pada analisis 10 butir soal pilihan ganda pada elemen akuntansi keuangan kelas xi SMK ini menghasilkan 6 soal di nomor 1, 2, 4, 5, 8, dan 9 dengan koefisien korelasi yang signifikan dan sangat signifikan dengan skor total. Selanjutnya, 4 soal pada nomor 3,6,7, dan 10 yang mendapatkan koefisien korelasi yang tidak signifikan. Jika butir soal tersebut signifikan atau sangat signifikan dapat dianggap valid dan berkontribusi pada pengukuran konstruk yang diinginkan, sehingga butir soal bisa dipertahankan. Tapi untuk item soal tidak signifikan perlu dianalisis lebih mendalam, dengan demikian, item soal dapat dilakukan revisi atau harus dihapus atau dibuang (Hamimi, Zamharirah, & Rusydy, 2020)



No Butir Baru	No Butir Asli	Korelasi	Signifikansi
1	1	0,856	Sangat Signifikan
2	2	0,847	Sangat Signifikan
3	3	0,498	-
4	4	0,699	Signifikan
5	5	0,699	Signifikan
6	6	0,036	-
7	7	0,500	-
8	8	0,683	Signifikan
9	9	0,578	Signifikan
10	10	0,345	-

**Gambar.3 Hasil Validitas**

### Daya Pembeda

Potensi instrumen tes dalam mengenali variasi kompetensi peserta atau membedakan tingkat keahlian, prestasi, atau kualitas peserta tes. Selain mampu mengidentifikasi variasi kompetensi yang dimiliki peserta tes daya pembeda juga dapat berfungsi untuk membantu menilai kualitas dan efektivitas sebuah tes. Daya pembeda dapat digunakan untuk memastikan bahwa soal-soal dalam tes mampu mengukur variasi kemampuan dengan akurat, sehingga dapat memberikan informasi yang valid tentang

tingkat penguasaan materi para peserta tes (Fiska et al., 2021). Menggunakan teknik statistik koefisien korelasi biserial dalam pengelolaan data untuk menganalisis korelasi dengan membandingkan skor individu soal terhadap total nilai peserta. Kisaran koefisien pembeda soal terletak antara -1,00 hingga 1,00, tanda negatif pada butir soal menentukan soal tersebut tidak valid karena peserta kelompok pandai gagal menjawab, sedangkan kelompok kurang pandai berhasil menjawab benar (Nurhalimah, Hidayati, Rosidi, Wiwin, & Hadi, 2022). Kriteria untuk mengkategorikan daya pembeda sebagai berikut :

**Tabel 2. Kriteria Indeks Daya Beda**

Indeks daya beda	Kategori
0,00 - 0,20	Jelek (Soal diperbaiki)
0,21 - 0,40	Cukup (Soal diterima dan diperbaiki)
0,41 - 0,70	Baik (Soal diterima)
0,71 - 1,00	Sangat baik (Soal diterima)
Negatif	Semuanya tidak baik (Soal dibuang)

Daya pembeda bernilai 0 ketika performa menjawab yang serupa antara kelompok berkemampuan tinggi dan rendah. Interval skor pembeda dari -1 hingga +1, dengan fokus pada nilai positif. Dimana mendekati 1 menunjukkan potensi butir soal dalam mengidentifikasi variasi keahlian peserta secara lebih baik. Butir soal dengan kapasitas pembeda yang sangat efektif adalah item yang lebih dari 0,5. Hasil evaluasi kapasitas pembeda pada instrumen tes pilihan ganda elemen akuntansi keuangan kelas xi SMK yang menggunakan *software* anates V4 terlihat menunjukkan 8 dari 10 soal memperlihatkan daya pembeda di atas 0,5, terletak dinomor 1, 2, 3, 4, 5, dan 9 yang dinyatakan sangat baik dan nomor 8 dan 10 dinyatakan baik, artinya butir soal ini memiliki kualitas yang tinggi dalam memisahkan peserta berdasarkan tingkat penguasaan materi. Selanjutnya 2 dari 10 soal dengan daya beda rendah dari 0,5, berlokasi pada nomor 6 dan 7 yang dinyatakan cukup. Perlu analisis lebih lanjut untuk butir soal yang

dinyatakan cukup, supaya dapat diputuskan bahwa soal tersebut dapat digunakan atau diperbaiki.

No Butir Baru	No Butir Asli	Kel. Atas	Kel. Bawah	Beda	Indeks DP (%)
1	1	5	0	5	100,00
2	2	5	0	5	100,00
3	3	5	1	4	80,00
4	4	5	1	4	80,00
5	5	5	1	4	80,00
6	6	2	0	2	40,00
7	7	5	3	2	40,00
8	8	4	1	3	60,00
9	9	5	1	4	80,00
10	10	4	1	3	60,00

**Gambar.4 Daya Beda**

### Tingkat Kesukaran

Tingkat kesukaran merupakan parameter penting dalam analisis kualitas instrumen tes yang menggambarkan seberapa sulit atau mudahnya suatu butir soal dipahami dan dijawab oleh peserta didik (Kaka, Baro, & Njoeroemana, 2024). Konsep ini bertujuan untuk mengidentifikasi rasio peserta yang dapat menjawab setiap soal dengan tepat, sehingga memberikan hasil tes yang berimbang dan representatif. Indeks kesukaran 0,00 menunjukkan bahwa soal sangat sulit dimana seluruh siswa gagal menjawab soal dengan akurat, sementara skor sempurna yaitu 1,00 mengungkapkan bahwa soal sangat mudah karena dapat diselesaikan dengan teoat oleh semua siswa (Akhmadi, 2021). Secara teknis, tingkat kesukaran diukur melalui rasio keberhasilan peserta dalam menjawab soal dibandingkan keseluruhan peserta (Merlin Purwati, Arianty, Malaka Syakilah, Ridlo, & Susilaningsih, 2021) Nilai tingkat kesukaran bervariasi dari 0,00 sampai 1,00 (Elviana, 2020b), dengan kategori interpretasi sebagai berikut :

**Tabel 3. Indeks Tingkat Kesukaran**

Indeks Kesukaran	Kategori
$0,00 \leq TK \leq 0,30$	Sukar

$0,31 \leq TK \leq 0,70$       Sedang

$0,71 \leq TK \leq 1,00$       Mudah

---

Dari analisis item soal pada tes objektif berbentuk pilihan ganda pada elemen akuntansi keuangan kelas xi SMK yang menggunakan analisis *software* anates v4 menunjukkan tingkat kesukaran soal mayoritas termasuk dalam tingkat kesulitan menengah. Dengan sejumlah 1 soal dengan tingkat kesulitan mudah pada indeks kesukaran 85,00%, 8 item soal yang tergolong dalam rentang kesukaran sedang dengan skala kesukaran 35,00-70,00%, dan 1 item soal yang tergolong dalam rentang kesukaran sangat sukar dengan skala kesukaran 15,00%. Soal dengan tingkat kesukaran yang terlalu tinggi dapat menimbulkan keputusasaan pada para peserta tes, dengan keberadaan soal yang memiliki tingkat kesulitan sangat minim kurang mampu mengembangkan kemampuan peserta didik dalam menyelesaikan persoalan kompleks (Nurhalimah et al., 2022).

No Butir Baru	No Butir Asli	Jml Betul	Tkt. Kesukaran(%)	Tafsiran
1	1	8	40,00	Sedang
2	2	9	45,00	Sedang
3	3	14	70,00	Sedang
4	4	10	50,00	Sedang
5	5	10	50,00	Sedang
6	6	3	15,00	Sangat Sukar
7	7	17	85,00	Mudah
8	8	7	35,00	Sedang
9	9	14	70,00	Sedang
10	10	12	60,00	Sedang

**Gambar.5 Tingkat Kesukaran**

Melalui penggunaan Software Anates V4 untuk menganalisis soal pilihan ganda pada elemen akuntansi keuangan siswa kelas XI SMK, penelitian ini mengungkapkan metode strategis dalam mengembangkan instrumen evaluasi yang berkualitas. Proses optimasi dilakukan dengan mengevaluasi empat parameter kunci pengujian, yakni

reliabilitas, validitas, kemampuan membedakan tingkat pemahaman siswa, serta tingkat kualitas soal.

Proses optimasi menghasilkan temuan signifikan berdasarkan penelitian yang dilakukan. Pertama, koefisien reliabilitas tes mencapai 0,77, mengindikasikan tingkat kehandalan yang tinggi. Hal ini menunjukkan instrumen tes memiliki konsistensi dan keterpercayaan dalam mengukur kompetensi peserta didik pada elemen akuntansi keuangan.

Pada sisi validitas, penelitian mengindikasikan bahwa 6 dari 10 soal memiliki korelasi signifikan, sementara 4 soal memerlukan revisi atau penghapusan. Hasil analisis ini penting untuk memastikan setiap butir soal benar-benar mengukur konstruk yang dimaksud dengan akurat.

Hasil analisis pada daya pembeda mengungkapkan 8 soal berkualitas baik hingga sangat baik, dengan hanya 2 soal yang membutuhkan perbaikan. Hasil ini menunjukkan kapasitas instrumen dalam membedakan kemampuan peserta didik secara efektif.

Kemudian pada tingkat kesukaran soal bervariasi, dengan distribusi 1 soal mudah, 8 soal berkategori sedang, dan 1 soal sangat sukar. Variasi yang dihasilkan ini mencerminkan keseimbangan instrumen dalam menguji tingkat pemahaman peserta didik dari berbagai level kognitif.

Pendekatan sistematis menggunakan Anates V4 memungkinkan pendidik melakukan evaluasi mendalam, mengidentifikasi kelemahan instrumen, dan merancang strategis perbaikan yang tepat sasaran untuk meningkatkan kualitas penilaian pembelajaran akuntansi keuangan. Penelitian ini menegaskan pentingnya analisis secara komprehensif dalam pengembangan instrumen evaluasi, dengan Software Anates V4 sebagai alat bantu yang efektif dalam mengoptimalkan kualitas butir soal (Wati, Novita, & Iqbal Lubis, 2022).

## **KESIMPULAN DAN SARAN**

Pada penelitian analisis butir soal pilihan ganda pada elemen akuntansi keuangan kelas XI SMK guna mengoptimasi evaluasi pembelajaran dengan menggunakan Anates V4 ini dapat disimpulkan bahwa penelitian ini mengungkapkan metode strategis dalam mengembangkan instrumen evaluasi berkualitas dengan menggunakan Software Anates

V4. Fokus utama pada penelitian ini yaitu ada pada reliabilitas, validitas, daya pembeda, dan tingkat kesukaran.

Hasil analisis reliabilitas menggunakan Anates V4 menunjukkan koefisien realibilitas tes mencapai 0,77, menunjukkan tingkat kehandalan yang tinggi. Meskipun demikian, koefisien korelasi belum sepenuhnya dianggap handal dan membutuhkan evaluasi lebih mendalam. Validitas item memperlihatkan bahwa 6 dari 10 soal memiliki korelasi signifikan, sementara 4 soal memerlukan revisi atau penghapusan. Daya pembeda soal menunjukkan 8 soal berkualitas baik hingga sangat baik, dengan hanya 2 soal yang membutuhkan perbaikan. Tingkat kesukaran soal bervariasi, dengan 1 soal mudah, 8 soal berkategori sedang, dan 1 soal sangat sukar.

Pendekatan sistematis menggunakan Anates V4 membuktikan bahwa sebagai alat strategis bagi pendidik. Sehingga, pendidik dapat menganalisis secara mendalam pada item butir soal, penentuan kelemahan instrumen, dan mengembangkan strategi perbaikan kepada soal yang membutuhkan perbaikan yang tepat untuk meningkatkan evaluasi penilaian dalam pembelajaran akuntansi keuangan.

Para pendidik dan guru dianjurkan untuk memanfaatkan aplikasi Anates V4.0 dalam proses evaluasi secara berkala. Tujuannya adalah untuk melakukan penilaian mendalam terhadap kualitas soal yang dibuat, dengan harapan dapat menghasilkan instrumen penilaian yang lebih berkualitas. Melalui analisis yang komprehensif, diharapkan guru dapat merancang alat evaluasi yang tidak hanya valid dan handal, tetapi juga mampu menjawab keragaman kebutuhan dalam proses pembelajaran.

## **DAFTAR REFERENSI**

- Akhmadi, M. N. (2021). ANALISIS BUTIR SOAL EVALUASI TEMA 1 KELAS 4 SDN PLUMBUNGAN MENGGUNAKAN PROGRAM ANATES. *Ed-Humanistics : Jurnal Ilmu Pendidikan*, 6(1), 799–806. <https://doi.org/10.33752/ed-humanistics.v6i1.1464>
- Azzahroh, S., Iman, F. L., Anwar, B., & Aziz, R. (2022). Analisis butir soal ujian akhir semester mata kuliah psikologi belajar menggunakan software anates. *Journal of Indonesian Psychological Science (JIPS)*, 2(2). <https://doi.org/10.18860/jips.v3i2.17228>

- Butir, A., Evaluasi, S., & Akhmadi, M. N. (2021). *ANALISIS BUTIR SOAL EVALUASI TEMA 1 KELAS 4 SDN PLUMBUNGAN MENGGUNAKAN PROGRAM ANATES*.
- Drs. Syafril, M. d. (2019). *Dasar Dasar Ilmu Pendidikan* (1 ed.). Depok: Kencana.
- Elviana. (2020a). *ANALISIS BUTIR SOAL EVALUASI PEMBELAJARAN PENDIDIKAN AGAMA ISLAM MENGGUNAKAN PROGRAM ANATES* (Vol. 10).
- Eprillison, V., & Gumanti, D. (n.d.). *PENERAPAN APLIKASI ANATES DALAM MENGANALISIS BUTIR SOAL PADA GURU MATA PELAJARAN EKONOMI SMA YAPI PADANG*.
- Fakhri Ramadhan, M., Siroj, R. A., & Win Afgani, M. (2024). Validitas and Reliabilitas. *Journal on Education*, 06(02), 10967–10975.
- Fietri, W. A. (2021). *Natural Science: Jurnal Penelitian Bidang IPA dan Pendidikan IPA Analisis Butir Soal Biologi Kelas XI Madrasah Aliyah Sakinah Kerinci Menggunakan Program Komputer Anates 4.0 For Windows* (Vol. 7).
- Fietri, W. A., Lufri, Syamzurizal, & Zulyusri. (2021). *Analisis Butir Soal Biologi Kelas VIII Madrasah Tsanawiyah Negeri 6 Kerinci*.
- 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), 2021.
- Gusmizain, A. (2022). KARAKTERISTIK BUTIR SOAL TES MATA KULIAH MATRIKS & RUANG VEKTOR MAHASISWA MATEMATIKA. In *Jurnal Evaluasi Pendidikan* (Vol. 13).
- Hamimi, L., Zamharirah, R., & Rusydy. (2020). Analisis Butir Soal Ujian Matematika Kelas VII Semester Ganjil Tahun Pelajaran 2017/2018. In *MATHEMA JOURNAL E-ISSN* (Vol. 2).
- Ismanto. (2014). *EVALUASI\_HASIL\_BELAJAR\_PENDIDIKAN\_AGAMA\_ISLAM\_PAI*.

- Kaka, L., Baro, V. O., & Njoeroemana, Y. (2024). EFEKTIVITAS ANALISIS BUTIR SOAL PILIHAN GANDA MENGGUNAKAN APLIKASI ANATES DI SMPN 2 KANATANG. *Jurnal Inovasi Penelitian*.
- Khasanah, I., Fuady, A., & Sunismi. (2023). Analisis Soal Ulangan Harian Matematika Bentuk Pilihan Ganda. In *MATHEMA JOURNAL E-ISSN* (Vol. 5).
- Lasmy, L., Yusrizal, Y., & Razali, R. (2021). ANALISIS BUTIR SOAL USBN MATA PELAJARAN BAHASA INDONESIA DI SMA KABUPATEN ACEH BARAT. *Jurnal MUDARRISUNA: Media Kajian Pendidikan Agama Islam*, 11(3), 444. <https://doi.org/10.22373/jm.v11i3.5500>
- Magdalena, I., Fauziah, S. N., Faziah, S. N., & Nopus, F. S. (2021). ANALISIS VALIDITAS, RELIABILITAS, TINGKAT KESULITAN DAN DAYA BEDA BUTIR SOAL UJIAN AKHIR SEMESTER TEMA 7 KELAS III SDN KARET 1 SEPATAN. In *BINTANG : Jurnal Pendidikan dan Sains* (Vol. 3). Retrieved from <https://ejournal.stitpn.ac.id/index.php/bintang>
- Mania, S., Fitriani, F., 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). <https://doi.org/10.24252/asma.v2i2.16569>
- Mardiyah, H., & Rofi, H. (2024). *PENGGUNAAN APLIKASI ANATES PADA ANALISIS BUTIR SOAL PENILAIAN AKHIR JENJANG* (Vol. 11).
- Maulidah, H., Sukarno, S., & Syefrinando, B. (2023). ANALISIS KUALITAS INSTRUMEN TES FISIKA KELAS X MENGGUNAKAN SOFTWARE ANATES. *Physics and Science Education Journal (PSEJ)*. <https://doi.org/10.30631/psej.v2i3.1660>
- Merlin Purwati, L., Arianty, R., Malaka Syakilah, D., Ridlo, S., & Susilaningsih, E. (2021). *Analisis Soal Tes Pilihan Ganda Berbasis Higher Order Thinking Skill menggunakan Aplikasi Anates Windows Versi 4.0.9 For Windows*. Retrieved from [www.journal.uniga.ac.id](http://www.journal.uniga.ac.id)

- Muluki, A. (2020). Analisis Kualitas Butir Tes Semester Ganjil Mata Pelajaran IPA Kelas IV Mi Radhiatul Adawiyah. *Jurnal Ilmiah Sekolah Dasar*, 4(1). <https://doi.org/10.23887/jisd.v4i1.23335>
- Nurhalimah, S., Hidayati, Y., Rosidi, I., Wiwin, D., & Hadi, P. (2022). HUBUNGAN ANTARA VALIDITAS ITEM DENGAN DAYA PEMBEDA DAN TINGKAT KESUKARAN SOAL PILIHAN GANDA PAS. In *Jurnal Natural Science Educational Research* (Vol. 4).
- Prastika, Y. D. (2021). *PENGARUH VALIDITAS, RELIABILITAS DAN TINGKAT KESUKARAN TERHADAP KUALITAS BUTIR SOAL EKONOMI MENGGUNAKAN SOFTWARE ANATES DI SMKN 3 BANGKALAN*.
- Riani, D., Almujab, S., Dina, A., & Budiarto, R. (2020). *ANALISIS BUTIR SOAL DAN KEMAMPUAN SISWA DALAM MENJAWAB SOAL UJIAN NASIONAL PADA MATA PELAJARAN EKONOMI*.
- Sudijono, P. D. (2009). *Pengantar Statistik Pendidikan*. Raja Grafindo.
- Surapranata.Sumarna. (2006). *ANALISIS. VALIDITAS. RELIABILITAS DAN INTERPRETASI HASIL TES*. PT REMAJA ROSDAKARYA .
- Syamsurizal. (n.d.). *Validitas dan Reliabilitas Alat Ukur Syamsurizal*.
- Warju, W., Ariyanto, S. R., Soeryanto, S., & Trisna, R. A. (2020). ANALISIS KUALITAS BUTIR SOAL TIPE HOTS PADA KOMPETENSI SISTEM REM DI SEKOLAH MENENGAH KEJURUAN. *Jurnal Pendidikan Teknologi Dan Kejuruan*, 17(1). <https://doi.org/10.23887/jptk-undiksha.v17i1.22914>
- Wati, I., Novita, Y., & Iqbal Lubis, M. (2022). *Pelatihan Penggunaan Software Anatest dalam Menganalisis Butir Soal Tes Ekonomi bagi Mahasiswa Calon Guru*. 63(2), 2022. Retrieved from <https://ejournal.anotero.org/index.php/tasnim>