

## OPTIMALISASI KUALITAS SOAL PILIHAN GANDA KELAS 10 SMK AKUNTANSI MELALUI ANALISIS ANATES

Rebecca Avrely Serafim<sup>1\*</sup> Hastia Chandra Mustika Dewi<sup>2</sup> Luqman Hakim<sup>3</sup> Vivi Pratiwi<sup>4</sup>

<sup>1,2,3,4</sup>Program Studi Pendidikan Akuntansi, FEB, Universitas Negeri Surabaya, Indonesia

\*Email: rebbecavfim@gmail.com

### ARTICLE INFO

Received 7 December 2024

Revised 2 January 2025

Accepted 20 March 2025

Published 8 April 2025

#### Keywords:

anates, questions analysis, multiple choice

#### Kata Kunci:

anates, analisis butir soal, pilihan ganda

**To cite this article** Serafim, R., Dewi, H., Hakim, L., & Pratiwi, V. (2025). Optimalisasi Kualitas Soal Pilihan Ganda Kelas 10 SMK Akuntansi melalui Analisis Anates. *Jurnal Likhitaprajna*, 27(1), 28-38.  
<https://doi.org/10.37303/likhitaprajna.v27i1.360>



This is an open-access article under the [CC BY-SA](#) license.

Copyright (c) 2025 Rebecca Avrely Serafim, Hastia Chandra Mustika Dewi, Luqman Hakim, Vivi Pratiwi. Published by Fakultas Keguruan dan Ilmu Pendidikan Universitas Wisnuwardhana.

**Abstract:** This study aims to analyze the quality of test items used to measure students' abilities. Employing a descriptive quantitative approach with an evaluative method, the subjects of the study consisted of 23 test participants selected through purposive sampling. The object of the research was 10 multiple-choice questions with five answer options. Data were collected through documentation methods, including participants' answer sheets, answer keys, and total score data. The results show: (1) Test reliability is low (0.34); (2) Item validity varies with 30% moderate, 40% low, and 30% very low; (3) Item difficulty levels are disproportionate with 40% very easy, 50% easy, and 10% difficult; (4) Discrimination power shows 20% good items, 30% adequate, and 50% poor; (5) Distractor effectiveness reveals 40% non-functioning distractors, 45% minimally functioning, and 15% effectively functioning. Based on the analysis, the test instrument requires substantial improvement, particularly in reliability, difficulty level proportion, and distractor quality. Improvement recommendations include total revision for 30% of items, partial revision for 40% of items, and minor improvements for 30% of items.

**Abstrak:** Penelitian ini bertujuan untuk menganalisis kualitas butir soal dalam tes yang digunakan untuk mengukur kemampuan peserta didik. Menggunakan pendekatan deskriptif kuantitatif dengan metode evaluatif, subjek penelitian terdiri dari 23 peserta tes yang dipilih melalui teknik purposive sampling. Objek penelitian adalah 10 butir soal pilihan ganda dengan lima pilihan jawaban. Data dikumpulkan melalui metode dokumentasi, termasuk lembar jawaban peserta, kunci jawaban, dan data skor total. Hasil penelitian menunjukkan: (1) Reliabilitas tes tergolong rendah (0,34); (2) Validitas butir soal bervariasi dengan 30% validitas sedang, 40% rendah, dan 30% sangat rendah; (3) Tingkat kesukaran soal tidak proporsional dengan komposisi 40% sangat mudah, 50% mudah, dan 10% sukar; (4) Daya pembeda menunjukkan 20% soal berkategori baik, 30% cukup, dan 50% buruk; (5) Efektivitas pengecoh memperlihatkan 40% pengecoh tidak berfungsi, 45% berfungsi minimal, dan 15% berfungsi efektif. Berdasarkan hasil analisis, instrumen tes memerlukan perbaikan substansial terutama pada aspek reliabilitas, proporsi tingkat kesukaran, dan kualitas pengecoh. Rekomendasi perbaikan mencakup revisi total untuk 30% butir soal, revisi parsial untuk 40% butir soal, dan perbaikan minor untuk 30% butir soal.

## PENDAHULUAN

Pendidikan menjadi salah satu pilar paling utama dalam upaya pembangunan sumber daya manusia yang berkualitas. Di Indonesia, Sekolah Menengah Kejuruan (SMK) memiliki peranan penting dalam mempersiapkan lulusannya untuk memasuki dunia kerja dengan keterampilan yang relevan. Dalam konteks ini, evaluasi hasil belajar siswa menjadi aspek yang sangat krusial untuk memastikan bahwa proses pembelajaran yang dilakukan efektif dan selaras

dengan tujuan pendidikan. Bentuk soal yang umum digunakan adalah soal pilihan ganda, dimana bentuk soal ini diakui memiliki kemudahan dalam penilaian dan kemampuannya untuk mengukur berbagai tingkat pemahaman siswa secara objektif (Ristiliana et al., 2022; Sudarto & Abd. Hafid, 2023). Namun, untuk memastikan bahwa soal-soal tersebut berkualitas, tentu diperlukan analisis yang mendalam terhadap butir soal yang digunakan.

Aplikasi ANATES merupakan program pengolah data yang dirancang untuk membantu guru dalam menganalisis kualitas butir soal, mencakup validitas, reliabilitas, tingkat kesukaran, dan daya pembeda dari setiap butir soal (Elfira et al., 2023; Pendidikan et al., 2019). Dengan menggunakan ANATES, para pendidik dapat memperoleh informasi yang akurat mengenai seberapa jauh soal yang telah disusun dapat mengukur kemampuan siswa. Penelitian sebelumnya menunjukkan bahwa penggunaan ANATES dalam analisis soal pilihan ganda dapat meningkatkan kualitas tes dan memberikan umpan balik yang konstruktif bagi pengembangan instrumen evaluasi (Adi et al., 2023). Oleh karena itu, sangat penting untuk mengeksplorasi bagaimana aplikasi ini dapat diimplementasikan dalam konteks evaluasi soal pilihan ganda di tingkat SMK, khususnya untuk kelas 10.

Pemilihan soal yang berkualitas sangat berpengaruh terhadap hasil belajar siswa. Kualitas soal yang dirancang dengan baik tidak hanya dapat mengukur pengetahuan siswa, tetapi juga dapat mendorong siswa untuk berpikir kritis dan analitis (Istiqomah & Fitrianawati, 2023). Pada konteks pendidikan yang semakin menitikberatkan pada pengembangan keterampilan berpikir tingkat tinggi, perlu dipastikan bahwa soal yang digunakan tidak hanya sekadar menguji hafalan, tetapi juga kemampuan siswa dalam mengaplikasikan pengetahuannya ke dalam situasi yang lebih kompleks (Sahlan et al., 2020). Dengan demikian, analisis yang dilakukan menggunakan ANATES diharapkan mampu memberikan wawasan yang lebih mendalam mengenai kualitas soal beserta dampaknya terhadap proses pembelajaran.

Daya pembeda merupakan salah satu aspek penting yang perlu diperhatikan dalam analisis soal pilihan ganda. Tingkat daya pembeda mengacu pada kemampuan suatu soal untuk membedakan antara siswa yang memiliki kemampuan tinggi dan rendah. Butir soal yang baik harus memiliki daya pembeda yang tinggi, sehingga dapat memberikan informasi yang relevan mengenai kemampuan siswa (Mania et al., 2020). Dalam hal ini, ANATES dapat membantu pendidik dalam mengidentifikasi soal-soal yang tidak efektif dan memberikan rekomendasi untuk perbaikan (Elfira et al., 2023). Dengan demikian, penelitian ini bertujuan untuk mengevaluasi kualitas soal pilihan ganda yang digunakan di kelas 10 SMK dengan pemanfaatan aplikasi ANATES sebagai alat analisis.

Melalui penelitian ini, diharapkan dapat ditemukan pola dan karakteristik dari soal pilihan ganda yang efektif dan tidak efektif, sekaligus memberikan rekomendasi untuk pengembangan soal yang lebih baik di masa mendatang. Dengan demikian, penelitian ini tidak hanya akan memberikan kontribusi terhadap peningkatan kualitas evaluasi di SMK, tetapi juga terhadap peningkatan hasil belajar siswa secara keseluruhan. Penelitian ini akan menjadi langkah awal dalam upaya memperkuat kualitas tes di lingkungan pendidikan, khususnya di tingkat Sekolah Menengah Kejuruan (SMK), dengan memanfaatkan teknologi yang ada.

## METODE

Penelitian ini menggunakan pendekatan deskriptif kuantitatif dengan metode evaluatif untuk menganalisis kualitas butir soal. Subjek penelitian terdiri dari 23 peserta tes yang dipilih menggunakan teknik purposive sampling, dengan pertimbangan kelengkapan data hasil ujian. Objek penelitian adalah instrumen tes yang terdiri dari 10 butir soal pilihan ganda, masing-masing dengan lima pilihan jawaban. Pengumpulan data dilakukan melalui metode dokumentasi, yang mencakup pengambilan data sekunder berupa lembar jawaban peserta tes, kunci jawaban, data identitas peserta, dan data skor total.

Instrumen yang digunakan dalam penelitian ini meliputi program ANATES versi 4.09 untuk analisis butir soal, lembar dokumentasi untuk pencatatan data, dan rubrik penilaian kualitas butir soal. Analisis dilakukan dengan menggunakan program ANATES untuk mengevaluasi beberapa aspek kualitas butir soal, termasuk validitas, reliabilitas, daya pembeda, dan kualitas pengecoh (distraktor). Hasil analisis ini diharapkan dapat memberikan gambaran yang jelas mengenai efektivitas instrumen tes dalam mengukur kemampuan peserta, serta memberikan rekomendasi untuk perbaikan di masa mendatang.

## HASIL DAN PEMBAHASAN

Evaluasi pembelajaran merupakan komponen penting dalam proses pendidikan untuk mengukur tingkat pencapaian kompetensi peserta didik. Salah satu instrumen evaluasi yang umum digunakan adalah tes pilihan ganda. Namun, untuk memastikan kualitas instrumen tes tersebut, diperlukan analisis butir soal yang komprehensif. Analisis butir soal menjadi tahapan krusial untuk mengetahui karakteristik setiap butir soal dari aspek validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Seiring perkembangan teknologi, proses analisis butir soal dapat dilakukan secara lebih efisien menggunakan aplikasi komputer, salah satunya adalah Anates.

### Analisis Reabilitas Tes

Reliabilitas tes dalam aplikasi Anates merupakan aspek penting dalam pengembangan instrumen penilaian agar lebih efektif dan valid. Reliabilitas bergantung pada konsistensi dan kestabilan hasil pengukuran dengan menggunakan instrumen tes. Interpretasi koefisien reliabilitas adalah sebagai berikut:

**Tabel 1. Kriteria Korelasi Koefisien**

0,00 - 0,20	Sangat Rendah
0,21 - 0,40	Rendah
0,41 - 0,70	Sedang
0,71 - 0,90	Tinggi
0,91 - 1,00	Sangat Tinggi

Nilai koefisien reliabilitas biasanya berkisar antara 0 hingga 1. Semakin tinggi nilai koefisien, semakin tinggi pula reliabilitas tes. Berdasarkan hasil analisis menggunakan aplikasi ANATES, diperoleh nilai reliabilitas tes berikut:

No.Urut	No. Subyek	Kode/Nama Subyek	Skor Ganjil	Skor Genap	Skor Total
1	1	Rosa Agustina...	3	2	5
2	2	Hurin Putri S...	5	5	10
3	3	Nicholas Anjani	5	5	10
4	4	Jessica Putri	5	3	8
5	5	Zahra Damariva	4	3	7
6	6	Joel Pratama	5	5	10
7	7	Alvaro Theodore	3	4	7
8	8	Clarissa Tamara	4	4	8
9	9	Lanadila Putri	4	4	8
10	10	Nando araihan	4	3	7
11	11	Alika nazira	4	3	7
12	12	Clarissa Everina	5	4	9
13	13	Larasanaya Pu...	5	3	8
14	14	aurin nesta	5	4	9
15	15	Calista abara	4	3	7
16	16	Siti Fadilah	3	4	7
17	17	Febianti putri	4	4	8
18	18	Dian Rahmadani	5	2	7
19	19	Agniya Aprilia	4	4	8
20	20	Zidane Jiwandana	5	3	8
21	21	Agnia Nur Savina	3	3	6
22	22	Pratiwi sheena	5	3	8
23	23	Suryana	4	3	7

**Gambar 1. Hasil Reabilitas Tes**

Berdasarkan gambar diatas, diperoleh nilai reliabilitas tes sebesar 0,34. Dengan mempertimbangkan hasil analisis reliabilitas yang diperoleh dari aplikasi ANATES, dapat disimpulkan bahwa nilai rata-rata skor peserta tes adalah 7,78, yang menunjukkan rata-rata nilai yang diperoleh. Simpangan baku sebesar 1,24 mencerminkan tingkat variasi atau penyebaran skor di antara peserta tes. Selain itu, korelasi antara skor butir soal dan skor total adalah 0,21, yang menunjukkan adanya hubungan rendah. Nilai reliabilitas tes yang sebesar 0,34 menunjukkan bahwa tingkat keandalan atau konsistensi tes secara keseluruhan tergolong rendah. Dengan kata lain, tes yang disusun kurang dapat diandalkan untuk mengukur kemampuan peserta tes secara akurat. Hasil analisis ini secara keseluruhan menyimpulkan bahwa tes yang disusun memiliki reliabilitas yang rendah. Oleh karena itu, perlu dilakukan perbaikan dalam perumusan butir soal agar dapat menghasilkan tes yang lebih reliabel dan konsisten dalam mengukur kemampuan peserta tes.

### Analisis Kelompok Unggul

Analisis kelompok unggul dalam konteks pendidikan dan evaluasi pembelajaran merupakan langkah penting untuk memahami karakteristik dan kinerja siswa yang berada di atas rata-rata. Dalam hal ini, aplikasi Anates dapat digunakan sebagai alat untuk menganalisis dan mengevaluasi kualitas tes yang dirancang untuk mengukur kemampuan siswa, termasuk kelompok unggul.

Kelompok Unggul							
Nama berkas: C:\USERS\HP\DOCUMENTS\EVAL BELAJAR\ANALISIS_KELOMPOK.ANA							
No.Urut	No Subyek	Kode/Nama Subyek	Skor	1	2	3	4
1	2	Hurin Putri S...	10	1	1	1	1
2	3	Nicholas Anjani	10	1	1	1	1
3	6	Joel Pratama	10	1	1	1	1
4	12	Clarissa Everina	9	1	1	1	1
5	14	aurin nesta	9	1	1	1	1
6	4	Jessica Putri	8	1	1	1	1
Jml Jwb Benar				6	6	3	6
No.Urut	No Subyek	Kode/Nama Subyek	Skor	8	9	10	
1	2	Hurin Putri S...	10	1	1	1	
2	3	Nicholas Anjani	10	1	1	1	
3	6	Joel Pratama	10	1	1	1	
4	12	Clarissa Everina	9	1	1	1	
5	14	aurin nesta	9	1	1	1	
6	4	Jessica Putri	8	1	1	1	
Jml Jwb Benar				6	6	6	

Gambar 2. Hasil Kelompok Unggul

Berdasarkan informasi yang diperoleh dari gambar di atas, dapat disimpulkan bahwa kelompok unggul terdiri dari beberapa peserta yang berhasil meraih skor keseluruhan yang tinggi. Terdapat tiga peserta yang mencatatkan skor maksimal, yaitu 10, yang menunjukkan bahwa mereka memiliki penguasaan materi yang sangat baik dalam menjawab soal-soal tes. Selain itu, satu peserta lainnya berhasil mendapatkan skor 9, yang meskipun sedikit di bawah tiga peserta dengan skor 10, tetap menunjukkan kemampuan yang tinggi. Selanjutnya, terdapat satu peserta yang memperoleh skor 8, yang masih termasuk dalam kategori unggul meskipun berada sedikit di bawah peserta dengan skor 9.

Di samping itu, terdapat satu baris yang menunjukkan jumlah jawaban benar, dengan skor 6. Hal ini mengindikasikan bahwa jumlah skor dari beberapa peserta yang termasuk dalam kelompok unggul. Secara keseluruhan, kelompok unggul ini terdiri dari peserta-peserta yang mampu menguasai materi tes dengan baik, yang terbukti dari perolehan skor yang tinggi. Hal ini menunjukkan bahwa mereka memang memiliki kemampuan yang menonjol dalam bidang

atau topik yang diujikan, sejalan dengan penelitian yang menunjukkan bahwa penguasaan konsep yang baik berkontribusi pada hasil belajar yang lebih baik.

### Analisis Kelompok Asor

Analisis kelompok asor dalam konteks pendidikan dan evaluasi pembelajaran adalah proses yang penting untuk memahami bagaimana siswa dengan kemampuan yang berbeda berinteraksi dengan instrumen tes yang digunakan. Dalam hal ini, aplikasi Anates dapat digunakan untuk menganalisis kualitas dan performa instrumen tes, termasuk dalam konteks kelompok asor, yang merujuk pada kelompok siswa dengan kemampuan yang bervariasi.

Kelompok Asor							
Nama berkas: C:\USERS\HP\DOCUMENTS\EVAL BELAJAR\ANALISIS_KELOMPOK.ANA							
No.Urut	No Subyek	Kode/Nama Subyek	Skor	1	2	3	4
1	15	Calista abara	7	1	1	1	-
2	16	Siti Fadilah	7	1	1	1	-
3	18	Dian Rahmadani	7	1	-	1	-
4	23	Suryana	7	1	-	1	-
5	21	Agnia Nur Savina	6	1	1	-	-
6	1	Rosa Agustina...	5	-	-	1	1
Jml Jwb Benar				5	3	4	1
				5	5	5	5
No.Urut							
No Subyek							
1	15	Calista abara	7	-	-	1	
2	16	Siti Fadilah	7	1	-	1	
3	18	Dian Rahmadani	7	1	1	-	
4	23	Suryana	7	1	1	1	
5	21	Agnia Nur Savina	6	-	-	1	
6	1	Rosa Agustina...	5	1	1	-	
Jml Jwb Benar				4	3	4	

Gambar 3. Hasil Kelompok Asor

Pada kelompok asor, terdapat beberapa peserta yang memperoleh skor total yang lebih rendah dibandingkan dengan kelompok unggul. Dari hasil analisis yang ditunjukkan pada gambar di atas, terlihat bahwa terdapat lima peserta yang termasuk dalam kelompok asor. Di antara peserta tersebut, dua peserta berhasil meraih skor keseluruhan 7, yang tergolong lebih rendah dibandingkan dengan skor yang diperoleh peserta di kelompok unggul. Hal ini menunjukkan bahwa penguasaan materi tes yang dimiliki oleh kedua peserta tersebut masih kurang optimal. Selanjutnya, terdapat satu peserta yang memperoleh skor 6, yang juga lebih rendah dibandingkan dengan dua peserta lainnya yang mendapatkan skor 7. Skor ini mencerminkan bahwa penguasaan materi tes oleh peserta tersebut belum mencapai tingkat yang memuaskan.

Selain itu, satu peserta lainnya mencatatkan skor keseluruhan 5, yang merupakan skor terendah di antara seluruh peserta dalam kelompok asor. Ini menunjukkan bahwa peserta tersebut masih memiliki pemahaman yang cenderung kurang memadai terhadap materi tes. Secara keseluruhan, kelompok asor ini terdiri dari peserta-peserta yang menunjukkan tingkat penguasaan materi tes yang lebih rendah jika dibandingkan dengan kelompok unggul. Rata-rata hasil tes mereka berkisar antara 5 hingga 7, yang cukup jauh di bawah hasil yang diperoleh peserta di kelompok unggul. Kondisi ini mengindikasikan bahwa mereka masih memerlukan perhatian lebih agar dapat meningkatkan pemahaman dan penguasaan materi yang diujikan.

### Analisis Daya Pembeda

Analisis daya pembeda dalam konteks evaluasi pendidikan adalah proses penting untuk menentukan seberapa baik suatu butir soal dapat membedakan antara siswa yang memiliki

kemampuan tinggi dan rendah. Rentang nilai daya pembeda berkisar antara -1 hingga 1, dengan interpretasi sebagai berikut:

**Tabel 2. Kriteria Indeks Daya Beda**

0 - 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 merupakan tingkat kemampuan sebuah butir soal untuk dapat membedakan antara peserta yang menguasai materi (kelompok atas) dengan peserta yang kurang menguasai materi (kelompok bawah). Berdasarkan hasil analisis menggunakan aplikasi ANATES, diperoleh daya pembeda berikut:

DAYA PEMBEDA							
<hr/>							
Jumlah Subyek= 23							
Klp atas/bawah(n)= 6							
Butir Soal= 10							
Nama berkas: C:\USERS\HP\DOCUMENTS\EVAL BELAJAR\ANALISIS_KELOMPOK.ANA							
No	Butir Baru	No	Butir Asli	Kel. Atas	Kel. Bawah	Beda	Indeks DP (%)
1	1	1	6	5	1	16,67	
2	2	2	6	3	3	50,00	
3	3	3	6	4	2	33,33	
4	4	4	3	1	2	33,33	
5	5	5	6	5	1	16,67	
6	6	6	5	5	0	0,00	
7	7	7	6	5	1	16,67	
8	8	8	6	4	2	33,33	
9	9	9	6	3	3	50,00	
10	10	10	6	4	2	33,33	

**Gambar 4. Hasil Daya Pembeda**

Berdasarkan informasi yang disajikan, diketahui bahwa nilai daya pembeda untuk setiap butir soal adalah 6, yang berarti daya pembeda untuk semua butir soal adalah sama. Namun, nilai daya pembeda tidak dapat lebih dari 1. Oleh karena itu, pernyataan ini tidak akurat. Jika nilai daya pembeda keseluruhan butir soal menunjukkan nilai yang terletak dalam rentang 0,41 - 0,70, hal ini mengindikasikan bahwa butir soal memiliki daya pembeda yang cukup baik.

### Analisis Tingkat Kesukaran

Tingkat kesukaran soal adalah salah satu aspek penting dalam evaluasi kualitas butir soal, yang merujuk pada persentase jumlah siswa yang menjawab soal dengan benar. dengan interpretasi sebagai berikut:

**Tabel 3. Indeks Tingkat Kesukaran**

$0,00 \leq TK \leq 0,30$	Sukar
$0,31 \leq TK \leq 0,70$	Sedang
$0,71 \leq TK \leq 1,00$	Mudah

Indeks tingkat kesukaran biasanya dinyatakan dalam bentuk proporsi yang berkisar antara 0,00 hingga 1,00, di mana nilai 0,00 menunjukkan bahwa tidak ada siswa yang menjawab benar, dan nilai 1,00 menunjukkan bahwa semua siswa menjawab benar. Berdasarkan hasil analisis menggunakan aplikasi ANATES, diperoleh tingkat kesukaran berikut:

TINGKAT KESUKARAN							
=====							
Jumlah Subyek= 23							
Butir Soal= 10							
Nama berkas: C:\USERS\HP\DOCUMENTS\EVAL BELAJAR\ANALISIS_KELOMPOK.ANA							
No	Butir Baru	No	Butir Asli	Jml	Betul	Tkt. Kesukaran (%)	Tafsiran
1		1		20		86,96	Sangat Mudah
2		2		19		82,61	Mudah
3		3		20		86,96	Sangat Mudah
4		4		5		21,74	Sukar
5		5		17		73,91	Mudah
6		6		20		86,96	Sangat Mudah
7		7		22		95,65	Sangat Mudah
8		8		19		82,61	Mudah
9		9		19		82,61	Mudah
10		10		18		78,26	Mudah

**Gambar 5. Hasil Tingkat Kesukaran**

Dari analisis yang dilakukan, dapat disimpulkan bahwa setiap butir soal dalam tes ini memiliki nilai tingkat kesukaran yang berkisar antara 73,91% hingga 95,65%. Berdasarkan kriteria tingkat kesukaran, terdapat satu butir soal, yaitu soal nomor 5, yang termasuk dalam kategori sedang dengan nilai 73,91%. Sementara itu, butir soal nomor 1, 2, 3, 6, dan 7 tergolong mudah, dengan nilai yang diperoleh berkisar antara 86,96%. Butir soal nomor 8, 9, dan 10 juga termasuk dalam kategori mudah, dengan nilai 82,61%. Secara keseluruhan, hal ini menunjukkan bahwa tingkat kesukaran butir-butir soal dalam tes tersebut cenderung mudah. Hanya satu butir soal yang berada dalam kategori sukar, sedangkan mayoritas butir soal lainnya masuk dalam kategori mudah. Kondisi ini mengindikasikan bahwa tes ini mungkin kurang optimal dalam membedakan kemampuan peserta didik yang memiliki kemampuan tinggi dengan yang memiliki kemampuan rendah. Untuk meningkatkan kualitas tes, perlu dipertimbangkan untuk memasukkan butir-butir soal yang memiliki tingkat kesukaran yang lebih tinggi.

#### Analisis Korelasi Skor Butir dengan Skor Total

Analisis korelasi skor butir dengan skor total adalah langkah penting dalam evaluasi kualitas butir soal dalam instrumen tes. Korelasi ini memberikan informasi tentang seberapa baik setiap butir soal berkontribusi terhadap skor keseluruhan yang diperoleh oleh peserta tes. Nilai korelasi antara skor butir dan skor total (Korelasi) dapat diinterpretasikan sebagai berikut:

**Tabel 4. Kriteria Korelasi Validitas**

0,800 – 1,00	Validasi sangat tinggi
0,600 – 0,800	Validasi tinggi
0,400 – 0,600	Validasi cukup
0,200 – 0,400	Validasi rendah
0,00 – 0,200	Validasi sangat rendah

Pada konteks analisis butir soal, korelasi skor butir terhadap skor total bertujuan untuk mengukur seberapa kuat hubungan antara jawaban benar atas suatu butir soal dengan skor keseluruhan yang dicapai oleh peserta tes. Semakin tinggi korelasi, semakin baik butir soal tersebut dalam mengukur variabel tes secara keseluruhan. Berdasarkan hasil analisis menggunakan aplikasi ANATES, diperoleh korelasi berikut:

KORELASI SKOR BUTIR DG SKOR TOTAL				
<hr/>				
Jumlah Subyek= 23				
Butir Soal= 10				
Nama berkas: C:\USERS\HP\DOCUMENTS\EVAL BELAJAR\ANALISIS_KELOMPOK.ANA				
No Butir Baru	No Butir Asli	Korelasi	Signifikansi	
1	1	0,462	-	
2	2	0,485	-	
3	3	0,569	-	
4	4	0,268	-	
5	5	0,057	-	
6	6	0,356	-	
7	7	0,137	-	
8	8	0,201	-	
9	9	0,390	-	
10	10	0,426	-	

**Gambar 6. Hasil Korelasi Antara Skor Butir Dan Skor Total**

Berdasarkan hasil analisis, nilai korelasi antara skor butir dan skor total berada dalam rentang 0,057 - 0,485. Dari data yang disajikan, terlihat bahwa nilai korelasi tersebut berkisar antara 0,057 hingga 0,485. Berdasarkan interpretasi kriteria yang ada, hanya butir soal nomor 2 yang memiliki nilai korelasi 0,485, yang termasuk dalam kategori validitas soal yang cukup. Sementara itu, butir soal lainnya, yaitu nomor 1, 3, 4, 5, 6, 7, 8, 9, dan 10, masing-masing memiliki nilai korelasi di bawah 0,400, yang berarti berada dalam kategori validitas soal yang rendah atau sangat rendah. Dengan demikian, dapat disimpulkan bahwa secara keseluruhan, semua butir soal dalam tes ini memiliki validitas yang rendah atau sangat rendah dalam mengukur kemampuan yang sama terhadap skor total. Hanya satu butir soal yang masuk dalam kategori validitas yang cukup. Hal ini mengindikasikan bahwa tes tersebut perlu ditinjau ulang untuk meningkatkan kualitas butir-butir soal agar lebih efektif dalam mengukur kemampuan peserta tes.

### Analisis Kualitas Pengecoh

Analisis kualitas pengecoh dalam instrumen tes adalah proses penting untuk mengevaluasi efektivitas pilihan jawaban yang salah dalam soal pilihan ganda. Pengecoh yang baik harus mampu menarik perhatian peserta tes yang kurang memahami materi, sehingga mereka tidak dapat menjawab soal dengan benar. Berikut merupakan keterangan kualitas pengecoh:

**Tabel 5. Keterangan Kualitas Pengecoh**

**	Kunci jawaban
++	Sangat baik
+	Baik
-	Kurang baik
--	Buruk
---	Sangat buruk

Selanjutnya dapat dilakukan analisis terhadap kualitas pengecoh (distraktor) yang terdapat dalam soal tes ini. Mengacu pada kriteria yang ada, pengecoh dikategorikan baik jika dipilih oleh setidaknya 5% peserta tes, sedangkan pengecoh dianggap sangat baik jika dipilih oleh minimal 10% peserta tes. Pengecoh yang tidak berfungsi sama sekali adalah yang tidak dipilih oleh peserta tes. Selain itu, pengecoh yang berfungsi dengan baik akan menghasilkan nilai yang hampir setara dengan kunci jawaban, sedangkan pengecoh yang terlalu mudah akan menghasilkan nilai yang lebih tinggi dari kunci jawaban, dan pengecoh yang terlalu sulit akan menghasilkan nilai yang jauh lebih rendah dari kunci jawaban.

KUALITAS PENGECHOH									
=====									
Jumlah Subyek= 23									
Butir Soal= 10									
Nama berkas: C:\USERS\HP\DOCUMENTS\EVAL BELAJAR\ANALISIS_KELOMPOK.ANA									
No	Butir Baru	No	Butir Asli	a	b	c	d	e	*
1	1	1	2---	20**	0--	1+	0--	0	
2	2	2	2--	0--	1++	19**	1++	0	
3	3	3	0--	20**	1+	1+	1+	0	
4	4	4	1--	15---	5**	0--	2-	0	
5	5	5	1+	4---	0--	17**	1+	0	
6	6	6	0--	0--	20**	1+	2---	0	
7	7	7	1---	0--	0--	0--	22**	0	
8	8	8	19**	2--	0--	0--	2--	0	
9	9	9	0--	1++	1++	2--	19**	0	
10	10	10	0--	1++	1++	18**	3---	0	

Gambar 7. Hasil Kualitas Pengecoh

Hasil analisis terhadap data yang disajikan menunjukkan bahwa secara umum, pengecoh dalam soal tes ini memiliki kualitas yang cukup baik. Terdapat beberapa pengecoh dengan kualitas sangat baik, yaitu yang dipilih oleh lebih dari 10% peserta tes. Namun, ada juga beberapa pengecoh yang tidak berfungsi sama sekali, yang hanya dipilih oleh 0% peserta tes. Sebagian besar nilai pengecoh juga hampir setara dengan kunci jawaban, yang menunjukkan bahwa pengecoh-pengecoh tersebut berfungsi dengan baik. Meskipun demikian, masih ada beberapa pengecoh yang dinilai terlalu mudah, dengan nilai yang lebih tinggi dari kunci jawaban. Secara keseluruhan, kualitas pengecoh dalam soal tes ini dapat dikatakan cukup baik, tetapi masih perlu ditinjau kembali, terutama untuk pengecoh yang terlalu mudah.

## SIMPULAN

Melalui hasil analisis yang dilakukan terhadap kualitas butir soal dalam tes ini, diperoleh kesimpulan bahwa secara keseluruhan butir soal memiliki validitas dan reliabilitas yang rendah. Hasil analisis menunjukkan bahwa sebagian besar butir soal tidak mampu membedakan antara peserta yang memiliki kemampuan tinggi dan rendah, dengan hanya satu butir soal yang masuk dalam kategori validitas cukup. Hal ini mengindikasikan bahwa tes yang disusun tidak cukup efektif dalam mengukur kemampuan peserta secara akurat. Selain itu, analisis terhadap daya pembeda dan kualitas pengecoh menunjukkan bahwa terlepas dari beberapa pengecoh yang berfungsi dengan baik, ternyata masih terdapat pengecoh yang tidak berfungsi sama sekali. Hal ini menunjukkan diperlukannya perbaikan dalam penyusunan butir soal agar dapat meningkatkan efektivitasnya.

Dengan demikian, penelitian ini menegaskan pentingnya evaluasi berkala terhadap kualitas butir soal guna memastikan bahwa instrumen evaluasi yang digunakan mampu memberikan hasil yang valid dan reliabel. Adapun saran yang dapat diberikan kepada para pengembang soal adalah disarankan untuk melakukan revisi terhadap butir-butir soal yang memiliki kualitas rendah, serta mempertimbangkan untuk menambahkan soal-soal yang lebih menantang guna meningkatkan daya pembeda. Selain itu, pelatihan bagi para pengajar dalam merancang soal yang berkualitas dalam rangka meningkatkan kemampuan dalam menyusun instrumen evaluasi yang efektif. Dengan langkah-langkah tersebut, besar harapan agar kualitas tes dapat ditingkatkan, sehingga secara keseluruhan dapat lebih baik dalam mengukur kemampuan peserta didik.

## DAFTAR PUSTAKA

Adi, Y., Bakhtiar, N., Vebrianto, R., & Yuliastrin, A. (2023). Development of Thematic Concept Understanding Assessment In Class V Science Learning Content. *Sainsmat* :

- Jurnal Ilmiah Ilmu Pengetahuan Alam*, 12(1), 91.  
<https://doi.org/10.35580/sainsmat121391652023>
- Baharu, M. (2022). *Minda Baharu, Volume 6, No 2 Desember, 2022*. 6(2), 254–264.
- Bukifan, D., Yuliati, L., & Handayanto, S. K. (2020). Penguasaan Konsep Siswa Pada Materi Termodinamika Dalam Pembelajaran Argument Driven Inquiry for STEM Education. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, 5(8), 1121.  
<https://doi.org/10.17977/jptpp.v5i8.13917>
- Elfira, I., Sari, P. N., Syamsurizal, S., & Lufri, L. (2023). Analisis Butir Soal Uts Projek Ipas Kelas X Smk N1 Banda Aceh Memanfaatkan Anates 4.0 for Windows Pc Program. *Jurnal IPA Terpadu*, 7(2), 156. <https://doi.org/10.35580/ipaterpadu.v7i2.48346>
- Irma, Z. U., Kusairi, S., & Yuliati, L. (2020). Penguasaan Konsep Siswa pada Materi Fluida Statis Dalam Pembelajaran STEM Disertai E-Formative Assessment. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, 5(6), 822.  
<https://doi.org/10.17977/jptpp.v5i6.13638>
- Istiqomah, & Fitrianawati, M. (2023). Pengembangan Instrumen Soal berbasis Keterampilan Berpikir Tingkat Tinggi. *Buletin Edukasi Indonesia*, 2(01), 18–23.  
<https://doi.org/10.56741/bei.v2i01.224>
- Kurniawan, N. A., Hidayah, N., & Rahman, D. H. (2021). Analisis Kemampuan Berpikir Kritis Siswa SMK. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, 6(3), 334.  
<https://doi.org/10.17977/jptpp.v6i3.14579>
- 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), 274.  
<https://doi.org/10.24252/asma.v2i2.16569>
- Marambaawang, D., Bano, V. O., & Hada Enda, R. R. (2023). Analisis Kualitas Butir Soal Penilaian Akhir Semester Gasal Tahun 2021/2022 Menggunakan Iteman Di Smp Negeri 1 Kambera. *Dharmas Education Journal (DE\_Journal)*, 4(1), 233–243.  
<https://doi.org/10.56667/dejournal.v4i1.961>
- Ningrum, W. A., Rahmawati, R. C., Minarti, I. B., & Mekar, R. M. (2023). Analisis Butir Soal Ulangan Harian Pembelajaran IPA Pada Kelas VIII Di SMPN 21 Semarang. *Educatio*, 18(1), 91–101. <https://doi.org/10.29408/edc.v18i1.18291>
- Pendidikan, J., Garut, U., Purwati, L. M., Arianty, R., Syakilah, D. M., Ridlo, S., Universitas, P., & Semarang, N. (2019). *Analisis Soal Tes Pilihan Ganda Berbasis Higher Order Thinking Skill menggunakan Aplikasi Anates Windows Versi. 2015*, 460–473.
- Putri Juliani, R., & Erita, S. (2023). Analisis Validitas dan Reliabilitas Instrumen Penilaian Kemampuan Berpikir Kritis dalam Konteks Sekolah Menengah. *JEID: Journal of Educational Integration and Development*, 3(3), 169–179.  
<https://doi.org/10.55868/jeid.v3i3.313>
- Putri, R. M., & Erita, S. (2023). Pengukuran kemampuan pemecahan soal bilangan berpangkat, logaritma dan bentuk akar: evaluasi validitas dan reabilitas soal ujian. *Journal of Educational Integration and Development*, 3(4), 245–255.
- Ristiliana, R., Indah Wati, Yulia Novita, & M. Iqbal Lubis. (2022). Pelatihan Penggunaan Software Anatest dalam Menganalisis Butir Soal Tes Ekonomi bagi Mahasiswa Calon Guru. *Tasnim Journal for Community Service*, 3(2), 63–69.  
<https://doi.org/10.55748/tasnim.v3i2.153>
- Rizky, K. (2023). *Analisis Validitas Dan Reliabilitas Percobaan Soal Tes Sumatif Siswa Kelas X Semester 2*. 1–6.
- Sahlan, N. L., A. Ghani, A. R., & Ernawati, E. (2020). Pengaruh Tipe Soal Dan Jenis Kelamin Terhadap Hasil Belajar Biologi Siswa Sekolah Menengah Atas. *Jurnal Amal Pendidikan*, 1(2), 106. <https://doi.org/10.36709/japend.v1i2.13054>
- Sudarto, S., & Abd. Hafid, A. K. (2023). 3 1,2,3. *Jurnal Pengabdian Kepada Masyarakat*,

- 3(5), 793–800. <http://bajangjournal.com/index.php/J-ABDI>
- Warsah, I. (2022). Evaluasi Pembelajaran ( Konsep . Fungsi dan Tujuan ). *Jurnal Kajian Pendidikan Islam*, 1, 190.
- YUSUF, F. W. (2024). Analisis Butir Soal Asesmen Sumatif Biologi Materi Perubahan Lingkungan Dengan Menggunakan Anates Pada Kelas X Sma. *LEARNING: Jurnal Inovasi Penelitian Pendidikan Dan Pembelajaran*, 4(1), 126–135. <https://doi.org/10.51878/learning.v4i1.2777>