

## PEMANFAATAN APLIKASI ANATES UNTUK MENILAI KUALITAS SOAL HOTS DALAM EVALUASI PEMBELAJARAN : ANALISIS PADA SISWA KELAS X AKUNTANSI DI SEKOLAH MENENGAH KEJURUAN

Fitrotul Aliyah<sup>1</sup>, Sandy Yudha Mahendra<sup>2</sup>, Maria Imelda Rani Rahayu<sup>3</sup>, Luqman Hakim<sup>4</sup>,  
Vivi Pratiwi<sup>5</sup>

[fitrotul.23147@mhs.unesa.ac.id](mailto:fitrotul.23147@mhs.unesa.ac.id)<sup>1</sup>, [sandy.23018@mhs.unesa.ac.id](mailto:sandy.23018@mhs.unesa.ac.id)<sup>2</sup>, [maria.23041@mhs.unesa.ac.id](mailto:maria.23041@mhs.unesa.ac.id)<sup>3</sup>,  
[luqmanhakim@unesa.ac.id](mailto:luqmanhakim@unesa.ac.id)<sup>4</sup>, [vivipratiwi@unesa.ac.id](mailto:vivipratiwi@unesa.ac.id)<sup>5</sup>

Universitas Negeri Surabaya

### ABSTRAK

Penelitian ini bertujuan untuk menganalisis kualitas soal berbasis Higher Order Thinking Skills (HOTS) dengan menggunakan aplikasi Anates pada siswa kelas X Akuntansi di SMKN. Sebanyak 15 soal pilihan ganda yang dirancang untuk mengasah kemampuan analisis, evaluasi, dan sintesis siswa diuji, dengan data yang diperoleh dari 20 siswa. Analisis difokuskan pada aspek reliabilitas, tingkat kesulitan, daya pembeda, efektivitas pengecoh, dan validitas. Hasil analisis menunjukkan bahwa reliabilitas tes mencapai 0,92, yang menandakan bahwa instrumen tersebut cukup konsisten. Namun, hanya 8 dari 15 soal yang memenuhi kriteria validitas signifikan, sementara sisanya memiliki korelasi rendah dengan skor total. Sebagian besar soal tergolong sangat mudah (>90%), sehingga tidak cukup menantang siswa untuk berpikir kritis. Daya pembeda soal cukup baik, yang menunjukkan kemampuannya untuk membedakan siswa dengan kemampuan tinggi dan rendah. Selain itu, banyak pengecoh yang tidak berfungsi dengan baik, dengan 62,67% pengecoh tidak dipilih oleh siswa. Berdasarkan hasil temuan ini, sebagian besar soal perlu direvisi agar sesuai dengan standar soal HOTS yang efektif. Penggunaan aplikasi Anates terbukti membantu dalam mengidentifikasi kelemahan soal secara sistematis dan memberikan arahan untuk perbaikan. Penelitian ini diharapkan dapat memberikan kontribusi dalam meningkatkan kualitas instrumen evaluasi pembelajaran, terutama dalam penerapan soal berbasis HOTS di sekolah menengah kejuruan.

**Kata Kunci:** Higher Order Thinking Skills (HOTS), Kualitas Soal, Aplikasi Anates.

### PENDAHULUAN

Manusia sebagai sumber daya memiliki peran yang krusial dalam pembangunan suatu bangsa. Untuk meningkatkan kualitas SDM, terdapat berbagai aspek yang perlu diperhatikan, meliputi aspek - aspek seperti sikap, mentalitas, perilaku, keterampilan, kecerdasan, kepercayaan, aturan, dan kondisi kesehatan. Pendidikan adalah salah satu upaya utama untuk meningkatkan kualitas sumber daya manusia, sesuai dengan tujuan bangsa Indonesia yang termuat dalam pembukaan UUD 1945, yaitu, untuk mencerdaskan kehidupan bangsa. Pendidikan tidak hanya penting untuk mengembangkan pengetahuan, tetapi juga untuk mengelola alam dan kehidupan secara bijaksana.

Sumber daya manusia yang unggul bergantung pada pendidikan. Dalam era Revolusi Industri 4.0, siswa harus memiliki kemampuan berpikir tingkat tinggi. Dalam proses evaluasi pembelajaran, soal yang berbasis pada High Order Thinking Skills (HOTS) digunakan untuk meningkatkan kemampuan ini. HOTS dimaksudkan untuk menguji kemampuan siswa dalam menganalisis, mengevaluasi, dan menemukan solusi untuk masalah tertentu. Kemampuan berpikir kritis dan analitis sangat penting dalam pembelajaran SMK, terutama di jurusan akuntansi. Hal ini disebabkan oleh fakta bahwa dunia kerja modern menuntut orang yang memiliki kemampuan berpikir logis, kreatif, dan mampu memecahkan masalah kompleks.

Meskipun demikian, penerapan soal HOTS sering kali menemui berbagai kendala, terutama dalam menjamin kualitas soal yang digunakan. Soal dengan kualitas yang kurang baik dapat menghambat tujuan evaluasi dan menyebabkan hasil yang tidak menggambarkan kemampuan siswa secara akurat. Oleh karena itu, penulis membutuhkan alat yang dapat membantu mengevaluasi dan menganalisis kualitas soal secara objektif, salah satunya adalah Aplikasi Anates. Aplikasi Anates merupakan perangkat lunak yang dirancang untuk menganalisis berbagai aspek kualitas soal, seperti validitas, reliabilitas, daya pembeda, dan tingkat kesulitan. Melalui aplikasi ini, penulis dapat memastikan soal yang diberikan kepada siswa memenuhi kriteria yang baik dan sesuai dengan karakteristik soal HOTS.

Dengan menggunakan Aplikasi Anates, penelitian ini menyelidiki penggunaan soal HOTS oleh siswa kelas X yang belajar akuntansi di SMK. Hasil penelitian ini diharapkan tidak hanya memberikan gambaran yang komprehensif dan berkualitas tentang soal yang digunakan, tetapi juga menawarkan saran untuk meningkatkan kemampuan berpikir kritis siswa. Selain itu, penelitian ini juga diharapkan dapat membantu pendidik membuat soal yang lebih sesuai dengan kebutuhan siswa saat ini.

## **METODOLOGI**

Model penelitian yang diterapkan dalam studi ini adalah metode kuantitatif yang menggunakan pendekatan deskriptif. Penelitian ini ditujukan pada siswa kelas X jurusan Akuntansi di Sekolah Menengah Kejuruan (SMK) dengan tujuan untuk menilai keefektifan soal HOTS (Higher Order Thinking Skills) dalam evaluasi pembelajaran yang sudah dibuat sebelumnya. Proses penelitian dimulai dengan penyusunan pertanyaan berupa pilihan ganda yang mengukur sejauh mana kemampuan siswa dalam berpikir tingkat tinggi sesuai standart HOTS, yang kemudian dikemas dalam bentuk digital dan dibagikan kepada siswa secara online. Kemudian jawaban siswa dari soal tersebut diuji cobakan dengan menggunakan aplikasi Anates untuk menganalisis kualitas butir soal berdasarkan jawaban siswa, termasuk ketepatan, konsistensi, level kesulitan, dan daya pembeda setiap soal. Penggunaan aplikasi Anates ini membantu dalam pemrosesan data secara efisien, memberikan hasil analisis yang akurat untuk setiap butir soal.

Pengumpulan data dilakukan melalui tes online yang diakses oleh siswa di kelas X jurusan Akuntansi. Setelah siswa menyelesaikan tes, data respons mereka diimpor ke aplikasi Anates, di mana setiap butir soal dianalisis berdasarkan standar yang telah ditentukan. Hasil dari aplikasi Anates kemudian diinterpretasikan untuk menilai sejauh mana soal-soal tersebut dapat memenuhi standar soal HOTS yang sesuai, terutama dari sisi kualitas soal dalam mendorong siswa berpikir kritis dan analitis. Analisis ini juga digunakan untuk mengidentifikasi butir soal yang perlu diperbaiki atau disesuaikan, sehingga hasil penelitian ini dapat berkontribusi dalam pengembangan soal evaluasi yang lebih efektif untuk meningkatkan kualitas pembelajaran di SMK.

## **HASIL DAN PEMBAHASAN**

Hasil penilaian materi Kesehatan, Keselamatan Kerja, dan Lingkungan Hidup (K3LH), yang terdiri dari 15 soal pilihan ganda dengan sampel 20 siswa, digunakan dalam penelitian ini. Aplikasi Anates digunakan untuk menganalisis item soal, yang mencakup reliabilitas, tingkat kesulitan, daya pembeda, kualitas pengecoh, validitas, dan korelasi. Untuk memberikan gambaran yang lebih baik, hasil analisis masing-masing soal dirangkum.

### **1. Reabilitas Tes**

Uji reliabilitas adalah serangkaian pengukuran yang menunjukkan konsistensi atau kestabilan, dengan hasil yang dapat dipercaya yang dihasilkan oleh pengukuran yang dilakukan berulang kali dengan alat ukur yang sama. Dalam penelitian ini, analisis

reliabilitas dilakukan menggunakan program Anates; nilai indeks reliabilitasnya dianggap tinggi jika nilainya lebih dari 0,70, dan nilainya kurang dari 0,70 dianggap rendah. Hasilnya dapat dilihat pada Tabel berikut

```

RELIABILITAS TES
=====

Rata2= 13,15
Simpang Baku= 2,41
KorelasiXY= 0,85
Reliabilitas Tes= 0,92
Nama berkas: C:\USERS\ACER\DOCUMENTS\ANALISIS SOAL PILIHAN GANDA.ANA

```

No.Urut	No. Subyek	Kode>Nama Subyek	Skor Ganjil	Skor Genap	Skor Total
1	1	Muhammad Aldi...	5	5	10
2	2	Grace Mardale...	8	7	15
3	3	Putri Mardale...	8	7	15
4	4	Arini Widya A...	8	7	15
5	5	Sapti Juliana...	8	7	15
6	6	Lisma Dila	7	6	13
7	7	Sarifudin	8	7	15
8	8	Saina Meilani	8	7	15
9	9	Dimas Maulana	8	7	15
10	10	Rasyid Iqbal ...	8	7	15
11	11	Agus Tri Wahyu	8	7	15
12	12	Renaldi Sopianur	8	7	15
13	13	Gledis Rinadz...	4	5	9
14	14	Saiful Irsyad	7	6	13
15	15	Semprol	5	4	9
16	16	Fenny Yoan Ve...	6	4	10
17	17	Sintia	7	6	13
18	18	Syeila Kila S...	8	7	15
19	19	Yesofina	7	5	12
20	20	Safira Qurrot...	4	5	9

Menurut kriteria, item tersebut dianggap sangat reliabel (sempurna), dengan nilai 0,92, lebih tinggi dari 0,70. Jika skor yang diamati dan skor yang sebenarnya memiliki korelasi yang tinggi, tes dianggap reliabel. Faktor-faktor ini dipengaruhi oleh karakteristik peserta ujian, tergantung pada kelompoknya.

## 2. Tingkat Kesukaran

Persentase jawaban yang benar, yang lebih rendah menunjukkan soal yang lebih sulit, dan persentase jawaban yang lebih tinggi menunjukkan soal yang lebih mudah. Soal-soal yang baik tidak terlalu mudah atau terlalu sulit untuk dijawab. Ada tiga kategori kesulitan soal: sangat mudah, mudah, dan sulit. Software ANATES digunakan untuk menganalisis tingkat kesulitan dalam penelitian ini, dan hasilnya dapat dilihat pada tabel berikut.

```

TINGKAT KESUKARAN
=====

Jumlah Subyek= 20
Butir Soal= 15
Nama berkas: C:\USERS\ACER\DOCUMENTS\ANALISIS SOAL PILIHAN GANDA.ANA

```

No	Butir Baru	No Butir Asli	Jml Betul	Tkt. Kesukaran (%)	Tafsiran
	1	1	19	95,00	Sangat Mudah
	2	2	18	90,00	Sangat Mudah
	3	3	19	95,00	Sangat Mudah
	4	4	17	85,00	Mudah
	5	5	17	85,00	Mudah
	6	6	19	95,00	Sangat Mudah
	7	7	18	90,00	Sangat Mudah
	8	8	18	90,00	Sangat Mudah
	9	9	19	95,00	Sangat Mudah
	10	10	18	90,00	Sangat Mudah
	11	11	16	80,00	Mudah
	12	12	17	85,00	Mudah
	13	13	16	80,00	Mudah
	14	14	16	80,00	Mudah
	15	15	16	80,00	Mudah

Berdasarkan hasil analisis, ditemukan bahwa dari 20 tidaklah sulit berdasarkan data.

sebagian besar soal dianggap memiliki tingkat kesulitan dari sedang hingga sangat mudah. Hasil menunjukkan 46,67% soal termasuk dalam kategori mudah dan 53,33% termasuk dalam kategori sangat mudah, dengan persentase tidak ada soal yang dianggap sulit.

### 3. Daya Pembeda

Kemampuan suatu soal dalam membedakan peserta didik dengan tingkat kecakapan yang tinggi dan rendah disebut dengan daya pembeda soal. Soal yang memiliki daya pembeda tinggi biasanya akan dijawab benar oleh peserta didik yang memiliki kecakapan tinggi, sedangkan peserta didik dengan kecakapan yang rendah umumnya menjawabnya salah. Untuk memudahkan dalam proses penelitian ini dapat dianalisis menggunakan tabel di bawah ini :

INDEKS DAYA PEMBEDA	KRITERIA
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)

Hasil analisis daya pembeda butir soal HOTS materi Keselamatan, Kesehatan kerja, dan Lingkungan Hidup pada kelas X Akuntansi Siswa SMK dengan menggunakan Anates ditunjukkan sebagai berikut :

```

DAYA PEMBEDA
=====

Jumlah Subyek= 20
Klp atas/bawah(n)= 5
Butir Soal= 15
Nama berkas: C:\USERS\ACER\DOCUMENTS\ANALISIS SOAL PILIHAN GANDA.ANA

No Butir Baru No Butir Asli Kel. Atas Kel. Bawah Beda Indeks DP (%)
1 1 5 4 1 20,00
2 2 5 3 2 40,00
3 3 5 4 1 20,00
4 4 5 3 2 40,00
5 5 5 2 3 60,00
6 6 5 4 1 20,00
7 7 5 3 2 40,00
8 8 5 3 2 40,00
9 9 5 4 1 20,00
10 10 5 5 0 0,00
11 11 5 2 3 60,00
12 12 5 3 2 40,00
13 13 5 2 3 60,00
14 14 5 2 3 60,00
15 15 5 3 2 40,00

```

Dari data tersebut dapat disimpulkan bahwa terdapat 5 soal dengan kategori Jelek, 6 soal dengan kategori Cukup, dan 4 soal dengan kategori Baik. Dalam penggunaan aplikasi Anates ini dapat menghitung indeks daya pembeda soal dengan cukup efektif dikarenakan hasil dari perhitungan tersebut tidak jauh beda dengan hasil menggunakan perhitungan manual. Langkah selanjutnya yaitu menyimpan/menggunakan soal dengan daya pembeda yang kategorinya baik karena kualitasnya sudah memenuhi standar. Sedangkan, soal dengan daya pembeda yang kategorinya jelek ataupun cukup perlu diperbaiki kualitasnya dan dapat diajukan kembali dalam tes berikutnya.

### 4. Kualitas Pengecoh

Efektivitas pengecoh yaitu pola yang menggambarkan cara mempengaruhi pemilihan jawaban dengan mempertimbangkan berbagai opsi jawaban yang tersedia di setiap soal. Tujuannya yaitu untuk membuat siswa bingung saat dirinya belum memahami materi dengan baik, sehingga mereka dapat terlihat dari siswa yang lainnya dengan tingkat pemahaman materi yang lebih tinggi.

Sebuah pengecoh dianggap efektif jika mampu menarik perhatian siswa dengan cara membuat siswa yang belum terlalu memahami materi pembelajaran merasa ragu sehingga terbesit dalam pemikirannya untuk memilih pengecoh tersebut sebagai pilihan jawaban dari pertanyaan yang benar. Efisiensi dari pengecoh pada tiap soal dievaluasi berdasarkan skala yang mencakup kategori dari sangat efektif hingga sangat tidak efektif..

- Pengecoh dikategorikan sangat efektif jika semua pengecoh bekerja dengan baik
- Pengecoh dikategorikan efektif jika tiga pengecoh bekerja dengan baik
- Pengecoh dikategorikan kurang efektif jika hanya dua pengecoh yang bekerja
- Pengecoh dikategorikan tidak efektif jika hanya satu pengecoh yang bekerja
- Pengecoh dikategorikan sangat tidak efektif jika tidak ada pengecoh yang bekerja sama sekali

Hasil analisis kualitas pengecoh soal HOTS materi Keselamatan, Kesehatan kerja, dan Lingkungan Hidup pada kelas X Jurusan Akuntansi Siswa SMK dengan menggunakan Anates ditunjukkan sebagai berikut :

```

KUALITAS PENGECOH
=====
Jumlah Subyek= 20
Butir Soal= 15
Nama berkas: C:\USERS\ACER\DOCUMENTS\ANALISIS SOAL PILIHAN GANDA.ANA

```

No Butir Baru	No Butir Asli	a	b	c	d	e	*
1	1	0--	0--	0--	1---	19**	0
2	2	0--	2---	18**	0--	0--	0
3	3	0--	1---	19**	0--	0--	0
4	4	3---	0--	17**	0--	0--	0
5	5	17**	1+	1+	1+	0--	0
6	6	19**	0--	0--	1---	0--	0
7	7	0--	0--	18**	1--	1--	0
8	8	1--	18**	1--	0--	0--	0
9	9	0--	19**	1---	0--	0--	0
10	10	0--	1--	18**	1--	0--	0
11	11	0--	16**	4---	0--	0--	0
12	12	1+	0--	1+	1+	17**	0
13	13	1++	16**	1++	0--	2--	0
14	14	1++	16**	1++	1++	1++	0
15	15	0--	3---	16**	0--	1++	0

Keterangan :

+ + : Sangat Baik

+ : Baik

- : Kurang Baik

-- : Buruk

--- : Sangat Buruk

Berdasarkan data pada tabel tersebut dapat disimpulkan bahwa kualitas pengecoh soal mayoritas dikatakan buruk atau tidak berfungsi dengan efektif tiap butir soalnya. Oleh karena itu, soal yang memiliki kualitas soal buruk perlu direvisi termasuk mengganti soal dengan tingkat kualitas yang lebih tinggi lagi agar soal dapat digunakan dengan efektif.

## 5. Keterkaitan Skor Butir dan Skor Total

Pada konteks aplikasi Anates, keterkaitan antara skor dari butir dan skor total yang didapat adalah sebuah analisis yang bertujuan untuk mengukur seberapa baik butir soal berkontribusi terhadap skor total dalam sebuah tes. Keterkaitan ini mengindikasikan adanya hubungan antara skor yang diperoleh siswa pada satu butir soal dengan total skor siswa pada keseluruhan tes.

Jika hubungan skor butir dan skor total tinggi, mengindikasikan bahwasanya soal tersebut relevan dan konsisten dengan tujuan penilaian keseluruhan, serta memiliki peran yang baik dalam mengukur kemampuan atau kompetensi yang ingin dinilai dalam tes. Sebaliknya, jika keterkaitan rendah atau negatif, butir soal tersebut dianggap kurang relevan atau mungkin tidak efektif dalam mendukung penilaian keseluruhan.

KORELASI SKOR BUTIR DG SKOR TOTAL

=====

Jumlah Subyek= 20

Butir Soal= 15

Nama berkas: C:\USERS\USER\DOCUMENTS\ANALISIS SOAL 15 PG.ANA

No Butir Baru	No Butir Asli	Korelasi	Signifikansi
1	1	0.409	-
2	2	0.522	Signifikan
3	3	0.309	-
4	4	0.505	Signifikan
5	5	0.688	Sangat Signifikan
6	6	0.409	-
7	7	0.522	Signifikan
8	8	0.522	Signifikan
9	9	0.409	-
10	10	-0.043	-
11	11	0.674	Sangat Signifikan
12	12	0.384	-
13	13	0.527	Signifikan
14	14	0.674	Sangat Signifikan
15	15	0.505	Signifikan

Berdasarkan tabel keterkaitan antara skor butir dan skor total dapat disimpulkan dan dinilai bahwa :

**A. Jumlah Subjek dan Batas Signifikan :**

Dengan jumlah subjek sebanyak 20 (oleh karena itu  $df = 18$ ), batas signifikansi untuk korelasi adalah 0,444 di tingkat kepercayaan 95% ( $p = 0,05$ ) dan 0,561 pada tingkat kepercayaan 99% ( $p = 0,01$ ). Artinya, butir soal dengan korelasi di atas 0,444 dianggap signifikan dan memiliki kontribusi yang cukup baik terhadap skor total.

**B. Butir Soal yang Signifikan :**

Butir soal pada nomor 2, 4, 5, 6, 7, 8, 11, 13, 14, dan 15 memiliki nilai korelasi di atas 0,444, sehingga dapat dikategorikan sebagai signifikan dalam mendukung skor total. Di antara butir soal yang signifikan, soal nomor 5, 11, dan 14 bahkan memiliki nilai korelasi yang sangat tinggi (masing-masing 0,688, 0,674, dan 0,674) dan diberi label "Sangat Signifikan." Hal ini menunjukkan Bahwa pertanyaan - pertanyaan tersebut sangat sesuai untuk menilai kemampuan siswa yang diharapkan dalam tes.

**C. Butir Soal yang Tidak Sesuai :**

Butir soal pada nomor 1, 3, 9, dan 12 memiliki nilai korelasi di bawah 0,444, sehingga dianggap tidak signifikan atau kurang mendukung skor total. Hal ini mengindikasikan bahwa soal-soal tersebut mungkin tidak relevan atau kurang efektif dalam mengukur keseluruhan kemampuan yang diinginkan.

Butir soal nomor 10 memiliki korelasi negatif (-0,043), yang berarti butir ini memiliki hubungan negatif dengan skor total dan mungkin perlu diubah atau diperbaiki karena dapat mengganggu konsistensi penilaian.

Ini berarti bahwa sebagian besar item pertanyaan dalam tes ini memiliki hubungan yang signifikan dengan total skor, yang menunjukkan kualitas dari butir pertanyaan ini yang cukup baik. Adapun sebagian butir pertanyaan yang tidak signifikan (khususnya butir 10 dengan korelasi negatif) sebaiknya diperbaiki atau dievaluasi ulang untuk memastikan

keseluruhan kualitas tes dalam mengukur kemampuan siswa dengan akurat.

## 6. Rekap Analisis Butir Soal

REKAP ANALISIS BUTIR  
=====

Rata2= 13.10  
Simpang Baku= 2.36  
KorelasiXY= 0.80  
Reliabilitas Tes= 0.89  
Butir Soal= 15  
Jumlah Subyek= 20  
Nama berkas: C:\USERS\USER\DOCUMENTS\ANALISIS SOAL 15 PG.ANA

Btr Baru	Btr Asli	D.Pembeda(%)	T. Kesukaran	Korelasi	Sign. Korelasi
1	1	20.00	Sangat Mudah	0.409	-
2	2	40.00	Sangat Mudah	0.522	Signifikan
3	3	20.00	Sangat Mudah	0.309	-
4	4	40.00	Mudah	0.505	Signifikan
5	5	60.00	Mudah	0.688	Sangat Signifikan
6	6	20.00	Sangat Mudah	0.409	-
7	7	40.00	Sangat Mudah	0.522	Signifikan
8	8	40.00	Sangat Mudah	0.522	Signifikan
9	9	20.00	Sangat Mudah	0.409	-
10	10	0.00	Mudah	-0.043	-
11	11	60.00	Mudah	0.674	Sangat Signifikan
12	12	40.00	Mudah	0.384	-
13	13	60.00	Mudah	0.527	Signifikan
14	14	60.00	Mudah	0.674	Sangat Signifikan
15	15	40.00	Mudah	0.505	Signifikan

### A. Daya Pembeda (D. Pembeda %):

Daya pembeda dalam tabel ini berkisar antara 0% hingga 60%. Nilai daya pembeda yang lebih tinggi (40%-60%) cenderung menunjukkan soal yang lebih baik dalam membedakan peserta dengan kemampuan yang berbeda. Ada satu soal (soal nomor 10) dengan daya pembeda sebesar 0%, dimana menunjukkan bahwa soal tersebut tidak dapat membedakan antara peserta yang memiliki kemampuan tinggi dan rendah, sehingga perlu dilakukan evaluasi kembali. Demikian juga dengan soal yang daya pembedanya di bawah 40% (misalnya soal nomor 1, 3, 6, dan 9) menunjukkan efektivitas pembeda yang lemah dan mungkin perlu direvisi untuk meningkatkan kualitasnya.

### B. Taraf Kesukaran (T. Kesukaran):

Soal di tabel tersebut sebagian besar masih ada pada golongan kategori "Sangat Mudah" dan "Mudah". Soal yang memiliki kategori "Sangat Mudah" mempunyai daya pembeda yang cukup rendah yaitu 20% - 40%. Sedangkan soal yang ada dalam kategori "Mudah" memiliki daya pembeda yang cukup berbeda yaitu 40% - 60%. Jika tes yang diberikan bertujuan untuk mengukur sejauh mana kemampuan peserta yang beragam, maka penambahan soal dengan taraf kesukaran standar atau sulit perlu untuk dipertimbangkan kembali mengingat untuk keseimbangan tingkat kesukaran.

### C. Korelasi dan Signifikansi Korelasi :

Korelasi antara skor butir dan total skor digunakan untuk mengukur validitas soal. Korelasi yang tinggi menunjukkan bahwa soal relevan dalam mengukur kemampuan yang diharapkan. Soal dengan korelasi signifikan dan sangat signifikan (misalnya soal dengan nomor 2, 4, 5, 7, 8, 11, 13, 14, 15) menunjukkan bahwa terdapat hubungan yang kuat pada skor butir dan skor keseluruhan, menandakan korelasi yang baik. Terdapat soal dengan korelasi rendah atau negatif (misalnya soal nomor 10 dengan korelasi -0.043), yang mengindikasikan bahwa soal ini mungkin tidak relevan dalam konteks tes dan perlu diperbaiki atau bahkan dihilangkan.

Sehingga dapat disimpulkan bahwa soal dengan daya pembeda tinggi (40%-60%) dan korelasi signifikan atau sangat signifikan (misalnya soal nomor 5, 11, dan 14) dianggap memiliki kualitas baik, karena mampu membedakan peserta dan relevan terhadap tujuan pengukuran. Sedangkan soal dengan daya pembeda rendah dan korelasi tidak signifikan atau

negatif (misalnya soal nomor 10 dan soal nomor 12) sebaiknya ditinjau kembali untuk perbaikan atau dihapus dari tes.

Secara keseluruhan, penambahan soal yang lebih kritis dan meningkatkan variasi kesulitan akan meningkatkan keandalan tes dalam mengukur beragam tingkat pengetahuan yang dimiliki peserta didik dalam memahami dan memaknai pengetahuan yang sebelumnya telah diberikan sebelum tes berlangsung.

## KESIMPULAN

Dari analisis yang telah dilakukan menggunakan media Anates diatas dapat diambil kesimpulan bahwasanya butir soal pilihan ganda berbentuk soal HOTS pada reliabilitas tes sebesar 0,92 ada dalam dalam klasifikasi baik. Tingkat kesulitan dari analisis tersebut yaitu mudah, dikarenakan mayoritas soal dijawab benar oleh peserta didik. Apabila diperhatikan berdasarkan daya pembeda soal masih ditemukan beberapa pertanyaan yang harus direvisi dikarenakan masih termasuk dalam kategori jelek ataupun cukup. Dengan hasil rekapitulasi 11 soal harus harus direvisi dan 4 soal dapat digunakan dengan baik.

Selain itu, masih terdapat beberapa pertanyaan yang mempunyai klasifikasi kesulitan yang kurang dari batas normal dan efektivitas dari pengecoh pertanyaan yang kurang optimal. Revisi difokuskan pada peningkatan validitas, reliabilitas, tingkat kesulitan, daya pembeda soal dan efektivitas pengecoh dari pertanyaan. Sedangkan untuk pertanyaan dengan kualitas normal dapat dipergunakan kembali untuk mengevaluasi kegiatan belajar mengajar agar target untuk memahami materi kepada peserta didik dapat tercapai.

## DAFTAR PUSTAKA

- Aliman, M., Mutia, T., Halek, D., Hasanah, R., & Muhammad, H. (2020). Pengembangan Instrumen Tes Kemampuan Berpikir Spasial Bagi Siswa SMA. *Geodika : Jurnal Kajian Ilmu dan Pendidikan Geografi*, 1 - 10.
- Amelia, N., Paridjo, P., & Sina, I. (2021). Analisis Butir Soal Penilaian Akhir Tahun Mata Pelajaran Matematika Menggunakan Software Anates. *AKSIOMA : Jurnal Matematika dan Pendidikan Matematika*, 447 - 456.
- Anarty, N., Suriyana, S., & Fitriana, I. (2024). Pengaruh Model Pembelajaran Reciprocal Teaching Terhadap Kemampuan Pemecahan Masalah Peserta Didik SMA Al- Ishlah Pontianak. *Cendikia : Jurnal Pendidikan dan Pengajaran*, 639 - 651.
- Andriani, L., & Handayani, R. Pemanfaatan Aplikasi Anates dalam Menganalisis Butir Soal untuk Meningkatkan Efektivitas Evaluasi Pembelajaran.
- Cahyaningrum, I., Fuady, A., & Sunismi, S. (2023). Analisis Soal Ulangan Harian Matematika bentuk Pilihan Ganda. *MATHEMA : Jurnal Pendidikan Matematika*, 110 - 125 .
- Fatimah, L. (2019). Analisis Kesukaran Soal, Daya Pembeda dan Fungsi Distraktor . *Jurnal Komunikasi dan Pendidikan Islam*, 37 - 64.
- Fitriyani, D., Putri, H., Andini, S., Cibro, D., & Pratiwi, A. (n.d.). Kemampuan Pemahaman Matematis Siswa Kelas II Sekolah Dasar Mengenai Operasi Hitung Bilangan Cacah . *Metodik Didaktik : Jurnal Pendidikan Ke - SD - an*, 1 - 14.
- Gusmizain, A. (2022). Karakteristik Butir Soal Tes Mata Kuliah Matriks & Ruang Vektor Mahasiswa Matematika. *Jurnal Evaluasi Pendidikan*, 127 - 130.
- Haq, A., & Anwar, N. (2024). Optimizing Arabic Language Assessment Through Anates Analysis. *Indonesian Journal of Islamic Studies* .
- Hartono, I., Tenriawaru, A., & Ningsih, K. (n.d.). Analisis Butir Soal Penilaian Sumatif IPA kelas VII SMP Negeri 3 Pontianak Menggunakan Anates . *Jurnal Kajian Pembelajaran dan Keilmuan*, 162 - 171.
- Herlina, T., & Saputra, M. (2018). Efisiensi dan Akurasi Penggunaan Anates untuk Analisis Butir Soal pada Evaluasi Berbasis Kompetensi. *Jurnal Ilmu Pendidikan*, 103 - 112.
- Hidinia, D., & Suprijono, A. (2020). Analisa Tingkat Kesulitan Soal Latihan USBN 2019 Melalui

- Aplikasi Anates. *Avatara : E-Journal Pendidikan Sejarah*, 1 - 5.
- Immanuel, C., Manik, S., Nababan, A., Sianturi, B., & Hasnah, A. (2024). Analisis Butir Soal Pilihan Ganda Mata Pelajaran Matematika Materi Bangun Data Berbasis Budaya Lokal di SDN 106163 Bandar Klipa. *Jurnal Ilmiah Multidisiplin Terpadu*.
- Kaka, L., Bano, V., & Njoeroemana, Y. (2024). Efektivitas Analisis Butir Soal Pilihan Ganda Menggunakan Aplikasi Anates di SMPN Kanatang. *Jurnal Inovasi Penelitian*, 1441 - 1450.
- Magdalena I, Fauziah, S., Faziha, S., & Nopus, F. (2021). Analisis Validitas, Reliabilitas, Tingkat Kesulitan dan Daya Beda Butir Soal Ujian Akhir Semester 7 Kelas III SDN Karet 1 Sepatan.
- Nur, A. (2020). Analisis Kualitas Butir Soal Penilaian Akhir Semester Gasal Mata Pelajaran Matematika Kelas IX SMP Negeri Ajibarang. *Doctoral Dissertation, IAIN Purwokerto*.
- Nurhalizah, S. (n.d.). Analisis Kualitas Butir Soal HOTS pada Mata Pelajaran Biologi yang Diberikan pada Siswa Kelas 12 dan Dievaluasi Menggunakan Media Anates.
- Pratama, H., & Santosa, E. (2021). Analisis Kualitas Butir Soal HOTS Menggunakan Aplikasi Anates pada Pembelajaran Berbasis Kompetensi. *Jurnal Pendidikan dan Pembelajaran*, 55 - 64.
- Purwati, L., Arianty, R., 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. *Jurnal Pendidikan UNIGA*, 460 - 473.
- Purwati, L. M., Arianty, R., 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. *Jurnal Pendidikan Universitas Garut*.
- Putra, A., & Yulianto, D. (2020). Penggunaan Aplikasi Anates dalam Analisis Soal HOTS di Sekolah Menengah Kejuruan. *Jurnal Pendidikan Matematika*, 43 - 50.
- Putri, R. A., & Rosliyah, Y. (2020). Analisis Distraktor Butir Soal Bunpou Shokyuu Kouhan Semester 2 Tahun 2018 / 2019 Prodi Pendidikan Bahasa Jepang Unnes. *Kiryoku*, 20 - 28.
- Ramadhani, D. (2024). Analisis Kualitas Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Informatika Kelas X di SMKN 8 Jember Tahun Ajaran 2022 / 2023. *Journal of Information Technology and Education (JITED)*, 14 - 26.
- Santosa, H. P. (2021). Analisis Kualitas Butir Soal HOTS Menggunakan Aplikasi Anates pada Pembelajaran Berbasis Kompetensi. *Jurnal Pendidikan dan Pembelajaran*, 55 - 64.
- Sari, V., Utomo, A., & Sumarwati, S. (2022). Kualitas Soal Bahasa Indonesia di SMP Muhammadiyah 1 Pontianak : Analisis Butir Soal. *Jurnal Pendidikan Bahasa dan Sastra Indonesia*, 112 - 119.
- Susanto, H., Rinaldi, A., & Novalia, N. (2015). Analisis Validitas Reliabilitas Tingkat Kesukaran dan Daya Beda Pada Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Matematika Kelas XII IPS di SMA Negeri 12 Bandar Lampung Tahun Ajaran 2014 / 2015. *Al Jabar : Jurnal Pendidikan Matematika*, 203 - 218.
- Tilaar, A., & Hasriyanti. (2019). Analisis Butir Soal Semester Ganjil Mata Pelajaran Matematika. *JP3I (Jurnal Pengukuran Psikologi dan Pendidikan Indonesia)*.
- Wulandari, S., & Rahmawati, L. (2020). Penggunaan Aplikasi Anates dalam Evaluasi Soal HOTS pada Mata Pelajaran Ekonomi. *Jurnal Evaluasi Pelajaran Ekonomi*, 123 - 134.
- Yusnita, N., & Fadilah, H. (2019). Evaluasi Soal Berbasis HOTS dengan Bantuan Aplikasi Anates di SMK. *Jurnal Teknologi dan Inovasi Pendidikan*, 97 - 108.
- Zarkasyi, Z., & Kartiko, A. (2024). Manajemen Kurikulum Satuan Pendidikan Mu'adalah. *Urwatul Wutsqo : Jurnal Studi Kependidikan dan Keislaman*, 168 - 178.