

PENERAPAN APLIKASI ANATES DALAM ANALISIS SOAL PILIHAN GANDA DI SMKN 2 TUBAN

**Nazwa Mirza Sabrina¹, Kirana Nayla Dewinta Hemi², Sri Hildayanti³, Lukman Hakim⁴,
Vivi Pratiwi⁵**

nazwa.23008@mhs.unesa.ac.id¹, kirana.23004@mhs.unesa.ac.id², sri.23098@mhs.unesa.ac.id³,
luqmanhakim@unesa.ac.id⁴, vivipratiwi@unesa.ac.id⁵

Universitas Negeri Surabaya

ABSTRAK

Penelitian ini bertujuan untuk mengetahui efektivitas menganalisis butir soal pilihan ganda dengan mata pelajaran Prinsip - Prinsip Akuntansi Dasar dan Perbankan Dasar Kelas X di SMKN 2 Tuban dengan menggunakan aplikasi Anates Versi 4.0. Penelitian ini merupakan penelitian deskriptif dengan pendekatan kuantitatif yang melibatkan 30 siswa sebagai sampel. Analisis dilakukan untuk mengukur validitas, reliabilitas, daya pembeda, tingkat kesukaran, kualitas pengecoh dan korelasi koefisien pada butir - butir soal. Penggunaan aplikasi Anates terbukti efektif dalam menganalisis kualitas soal, sehingga dapat digunakan sebagai alat untuk menyusun soal yang lebih berkualitas.

Kata Kunci: Analisis Butir Soal, Anates, Pilihan Ganda.

ABSTRACT

This research aims to determine the effectiveness of analyzing multiple choice questions in the subjects Basic Accounting Principles and Basic Banking in Class X at SMKN 2 Tuban using the Anates Version 4.0 application. This research is descriptive research with a quantitative approach involving 30 students as samples. Analysis was carried out to measure validity, reliability, differentiating power, level of difficulty, quality of distractors and correlation coefficients on the items. Using the Anates application has proven to be effective in analyzing the quality of questions, so it can be used as a tool to compose higher quality questions.

Keywords: Analysis Of Question Items, Anates, Multiple Choice.

PENDAHULUAN

Evaluasi merupakan proses yang dilakukan secara sistematis dengan tujuan untuk merencanakan, mengumpulkan, dan menyediakan informasi penting sebagai dasar pengambilan keputusan. Hal ini sejalan dengan pendapat Zainul dan Nasution (2001), menyatakan bahwa evaluasi merupakan keputusan yang didasarkan pada hasil pengukuran, yang bertujuan untuk menilai pencapaian hasil belajar siswa. Dalam proses pembelajaran, kegiatan evaluasi memiliki peran yang sangat penting karena berfungsi untuk mengukur sejauh mana tujuan pembelajaran tercapai. Melalui evaluasi, tingkat keberhasilan pengajar dalam melaksanakan tugas mengajar dapat dinilai berdasarkan kualitas metode yang digunakan, efektivitas strategi pengajaran, serta sejauh mana materi dapat dipahami oleh peserta didik. (Purwanto, 2011).

Dalam mengukur tingkat pemahaman siswa, pengajar akan membuat tes yang diujikan kepada siswa. Tes adalah alat yang digunakan untuk penilaian hasil belajar dan mengukur prestasi siswa sejalan dengan target penilaian. Tes terdiri dari serangkaian soal atau pertanyaan yang dirancang berdasarkan materi pembelajaran yang telah diajarkan. Jawaban siswa pada tes tersebut dinilai menggunakan skor atau skala yang telah ditetapkan, sehingga memberikan gambaran kuantitatif mengenai pencapaian siswa. Melalui tes, pengajar tidak hanya dapat mengukur prestasi siswa, tetapi juga mengevaluasi efektivitas metode pembelajaran yang digunakan. Dengan demikian, tes menjadi instrumen yang esensial dalam menilai keberhasilan proses pendidikan secara keseluruhan.

Agar tes dapat berfungsi secara optimal, diperlukan analisis untuk memastikan mutu dan efektivitasnya. Analisis tes merupakan salah satu langkah penting dalam mengkonstruksi tes yang bertujuan untuk memperoleh gambaran mengenai kualitas keseluruhan tes maupun kualitas setiap butir soal (Karno To, 2003). Analisis ini dilakukan setelah tes disusun dan dicobakan kepada peserta didik, dengan hasilnya dijadikan umpan balik untuk meningkatkan mutu tes. Proses analisis ini meliputi aspek kuantitatif dan kualitatif, yang bertujuan untuk mengetahui karakteristik setiap butir soal, seperti tingkat kesulitan, daya pembeda, serta keakuratan opsi jawaban.

Untuk mempermudah pendidik dalam mengevaluasi butir soal secara cepat dan praktis adalah menggunakan program komputer Anates versi 4.0 for Windows. Aplikasi ini membantu pendidik mengefisiensi waktu dalam melakukan standarisasi soal. Hasil analisis yang diperoleh melalui aplikasi ini dapat digunakan untuk merevisi, memperbaiki, atau bahkan mengganti soal yang kurang efektif.

Berdasarkan hasil uji coba soal yang dilakukan terhadap peserta didik kelas X di SMK Negeri 2 Tuban, diperoleh informasi bahwa dari 15 soal yang diujikan, sebanyak 11 soal dinyatakan layak digunakan, sedangkan 4 soal memerlukan revisi. Oleh karena itu, perlu dilakukan analisis lebih lanjut terhadap soal-soal tersebut untuk memastikan kualitas dan validitasnya.

1. Validitas Butir Soal

Validitas butir soal adalah sejauh mana suatu butir soal mendukung atau berkontribusi terhadap skor total yang dihasilkan dari keseluruhan tes. Validitas ini diukur dengan mengkorelasikan skor pada butir soal tertentu dengan skor total tes. Butir soal yang memiliki korelasi tinggi dianggap memiliki validitas yang baik karena mampu mencerminkan tujuan pengukuran tes secara keseluruhan.

Soal dengan korelasi tinggi menunjukkan relevansi yang kuat terhadap kompetensi yang ingin diukur, sehingga dianggap layak untuk digunakan kembali dalam tes berikutnya. Sebaliknya, butir soal dengan korelasi rendah dianggap kurang signifikan dan mungkin memerlukan revisi atau penggantian agar sesuai dengan standar kualitas evaluasi yang diharapkan. Validitas butir soal menjadi salah satu indikator utama dalam menentukan keefektifan dan keandalan sebuah tes.

2. Daya Pembeda Butir Soal

Daya pembeda soal adalah kemampuan suatu butir soal untuk membedakan antara peserta didik yang memiliki kemampuan tinggi dengan peserta didik yang memiliki kemampuan rendah. Indeks yang menggambarkan daya pembeda ini disebut indeks diskriminasi, dengan nilai berkisar antara 0,00 hingga 1,00. Indeks daya pembeda memberikan informasi tentang kualitas soal dalam menilai perbedaan pemahaman siswa terhadap materi yang telah diajarkan.

Butir soal dengan daya pembeda yang baik mampu mendeteksi sejauh mana peserta didik memahami materi. Sebaliknya, jika soal tidak mampu membedakan peserta didik berdasarkan tingkat pemahamannya, hal ini dapat disebabkan oleh beberapa faktor, seperti kunci jawaban yang tidak tepat atau adanya lebih dari satu jawaban yang benar pada butir soal tersebut.

Kriteria indeks daya pembeda adalah sebagai berikut:

- a. 0,00 – 0,20: Jelek (soal perlu diperbaiki).
- b. 0,20 – 0,40: Cukup (soal dapat diterima tetapi perlu revisi).
- c. 0,40 – 0,70: Baik (soal dapat diterima).
- d. 0,70 – 1,00: Sangat baik (soal dapat diterima).
- e. Negatif: Tidak baik (soal harus dibuang).

Melalui analisis daya pembeda, pendidik dapat meningkatkan mutu butir soal dengan memperbaiki, merevisi, atau mengganti soal yang tidak efektif. Proses ini memastikan bahwa soal yang digunakan dalam evaluasi mampu mengukur kemampuan siswa secara adil dan

akurat, serta memberikan gambaran yang jelas tentang pencapaian pembelajaran.

3. Tingkat Kesukaran Butir Soal

Tingkat kesukaran soal adalah ukuran yang menggambarkan sejauh mana suatu soal dapat dijawab dengan benar oleh peserta didik, yang dinyatakan dalam bentuk indeks dengan rentang nilai 0,00 hingga 1,00. Indeks tingkat kesukaran menunjukkan peluang peserta didik menjawab benar suatu soal pada tingkat kemampuan tertentu. Nilai $TK = 0,00$ berarti tidak ada siswa yang mampu menjawab soal tersebut dengan benar, sedangkan $TK = 1,00$ menunjukkan bahwa semua siswa menjawab soal tersebut dengan benar.

Semakin tinggi indeks tingkat kesukaran, semakin mudah soal tersebut, dan sebaliknya, semakin rendah indeksnya, semakin sulit soal tersebut. Soal yang baik memiliki tingkat kesukaran yang seimbang, yaitu tidak terlalu mudah dan tidak terlalu sulit, sehingga mampu memberikan pengukuran yang proporsional terhadap kemampuan siswa.

Tingkat kesukaran digunakan oleh pendidik untuk menganalisis hasil belajar siswa, mengevaluasi kesesuaian soal dengan kurikulum, dan mendeteksi adanya bias pada soal. Selain itu, tingkat kesukaran juga membantu dalam mengidentifikasi konsep yang perlu diajarkan ulang, mengevaluasi kelemahan dan kelebihan kurikulum, serta menyusun tes yang akurat dan sesuai. Soal dengan tingkat kesukaran yang tidak tepat mungkin memiliki kesalahan kunci jawaban, lebih dari satu jawaban benar, atau bentuk soal yang tidak relevan dengan materi yang diajarkan.

4. Reabilitas

Reliabilitas soal adalah tingkat ketepatan (precision) dan keajegan (consistency) suatu instrumen tes dalam menghasilkan skor yang stabil dan akurat. Indeks reliabilitas memiliki rentang nilai antara 0 hingga 1, di mana semakin tinggi nilai koefisien reliabilitas mendekati 1, semakin tinggi pula tingkat keajegan dan ketepatan tes tersebut. Reliabilitas menunjukkan sejauh mana suatu alat pengukur mampu secara konsisten dan handal mengukur apa yang dimaksudkan untuk diukur.

Tes yang memiliki reliabilitas tinggi akan memberikan hasil yang serupa jika digunakan dalam kondisi yang sama dan pada subjek yang sama. Dengan demikian, reliabilitas menjadi salah satu indikator utama untuk menilai kualitas suatu tes, karena tes yang andal dapat dijadikan acuan yang valid untuk mengevaluasi pencapaian peserta didik secara konsisten dan objektif.

METODE PENELITIAN

Penelitian ini merupakan penelitian deskriptif dengan pendekatan kuantitatif. Adapun tahapan yang dilakukan dalam penelitian ini adalah : 1) mengumpulkan hasil tes dari peserta didik SMKN 2 Tuban yang berupa jawaban dari soal pilihan ganda yang telah dikerjakan, 2) memasukkan hasil tes ke dalam aplikasi Anates 4.0, sehingga menghasilkan validitas, reliabilitas, daya pembeda, tingkat kesukaran, kualitas pengecoh, dan korelasi skor butir dengan skor total pada butir-butir soal, 3) analisis data dan pelaporan untuk mengolah data lalu dianalisa sehingga data dapat dibaca dan memiliki makna, kemudian dijadikan dasar penyusunan laporan sebagai media komunikasi peneliti dengan pembaca mengenai hasil penelitian.

HASIL DAN PEMBAHASAN

Hasil analisis ini digunakan untuk menilai kualitas butir soal berdasarkan data empiris yang diperoleh dari hasil tes. Analisis ini memungkinkan guru untuk mengambil langkah strategis berbasis data guna memperbaiki butir soal yang memiliki kelemahan, baik dari segi validitas, daya beda, maupun tingkat kesukaran. Upaya perbaikan tersebut tidak hanya bertujuan meningkatkan kualitas soal, tetapi juga memastikan bahwa evaluasi pembelajaran mampu mencerminkan kemampuan siswa secara lebih akurat, efektif, dan adil.

Proses analisis ini diharapkan menjadi bagian integral dari siklus pembelajaran, sehingga kualitas pendidikan dapat terus ditingkatkan secara berkelanjutan. Dengan demikian, penerapan metode ini tidak hanya berdampak pada peningkatan mutu soal, tetapi juga pada pengembangan kompetensi guru dalam merancang dan mengevaluasi instrumen pembelajaran secara profesional.

Berdasarkan data yang dianalisis, penelitian ini menggunakan soal pilihan ganda yang berfokus pada materi *Prinsip-Prinsip dan Konsep Akuntansi Dasar* serta *Perbankan Dasar* untuk kelas X di SMKN 2 Tuban. Sampel yang terlibat dalam analisis berjumlah 30 peserta didik, dengan jumlah soal sebanyak 15 butir dalam bentuk pilihan ganda. Hasil analisis butir soal mencakup sejumlah aspek penting, yaitu penghitungan reliabilitas tes, pengelompokan peserta didik menjadi kelompok unggul dan asor, analisis daya pembeda, tingkat kesukaran soal, korelasi skor butir dengan skor total, evaluasi kualitas pengecoh, yang selanjutnya digunakan untuk memberikan rekomendasi perbaikan atau penggunaan soal.

1. Reliabilitas Tes

Reliabilitas merujuk pada konsistensi dan kestabilan hasil yang dihasilkan oleh suatu alat ukur. Reliabilitas dihitung untuk menilai sejauh mana soal-soal yang disusun mampu memberikan hasil yang konsisten ketika digunakan dalam pengujian berulang. Tingkat reliabilitas yang baik menunjukkan bahwa instrumen pengukuran dapat dipercaya untuk menghasilkan hasil yang akurat.

Tabel 1 Hasil Reliabilitas

```

RELIABILITAS TES
*****
Rata-r: 10,87
Simpang Baku: 4,16
KorelasiKv: 0,85
Reliabilitas Tes: 0,92
Nama Berkas: BELUFI_001_NAMA_KSA

No. UJUT  No. Subyek  Kode/Nama Subyek  Skor Ganjil  Skor Genap  Skor Total
1         1         MOCHA ZAKARIA ROEP  2           1           3
2         2         ALPINA NURLIT...    6           4           10
3         3         MELLA KARISMA...   7           3           10
4         4         CLARA ALYSA...     5           4           9
5         5         ARINDA RIZKY ...   5           4           9
6         6         ADELVA ANTA P...   3           1           4
7         7         SITTE HADHAR...    8           7           15
8         8         DEVIKA SELVI H...  8           7           15
9         9         AULIA DEWI RA...   5           2           7
10        10        DINDA APRILIA...   7           6           13
11        11        DEVIN AULIA N...   6           6           12
12        12        ALFIYAN RUF...     4           6           10
13        13        DIMAS ANI DWI...   8           7           15
14        14        ANHAFI MEI KH...   7           6           13
15        15        DINA NUSITA R...   8           7           15
16        16        ANINDA KSA...      6           6           12
17        17        BINTANG ANGOR...   3           6           9
18        18        ANHISA DWI RY...   7           7           14
19        19        AISTE FERDINANTE   8           7           15
20        20        IIS WAPUNZ...      7           7           14
21        21        ACHMAD FERDIA...  8           7           15
22        22        LELYANA...         6           6           12
23        23        CANTIKA DWI E...   7           6           13
24        24        MELANY KUSUM...   1           1           2
25        25        ALDIAN HARDO...    2           0           2
26        26        DWI OKTAVIANA...  8           7           15
27        27        DESVITA APRIL...   8           7           15
28        28        NUSYA FITRI P...   5           1           6
29        29        YOGI HARDO PW...   7           7           14
30        30        TERRY HANUNG...    4           4           8
    
```

Berdasarkan data reliabilitas tes, diketahui bahwa rata-rata skor total yang diperoleh peserta didik adalah 10,87 dengan simpangan baku sebesar 4,16. Hal ini menunjukkan bahwa nilai rata-rata cukup baik, namun terdapat penyebaran skor yang menunjukkan variasi kemampuan peserta. Koefisien korelasi sebesar 0,85 mengindikasikan adanya hubungan yang sangat kuat antara skor ganjil dan skor genap, yang menunjukkan konsistensi internal soal. Selain itu, nilai reliabilitas tes sebesar 0,92 menunjukkan bahwa tes memiliki tingkat keandalan yang sangat tinggi. Tes dengan reliabilitas di atas 0,90 dianggap sangat baik dan dapat diandalkan untuk evaluasi pembelajaran. Namun, perlu dilakukan analisis lebih lanjut jika terdapat simpangan baku yang signifikan untuk menilai kesesuaian tingkat kesulitan soal terhadap kelompok siswa (Nurgiyantoro, 2013).

2. Kriteria Validitas

Korelasi antara skor butir dengan skor total merupakan indikator konsistensi internal suatu tes. Nilai korelasi menunjukkan sejauh mana sebuah butir soal berkontribusi terhadap pengukuran kemampuan yang dimaksudkan oleh keseluruhan tes (Arikunto, 2018). Nilai korelasi yang signifikan atau sangat signifikan menunjukkan bahwa butir soal tersebut memiliki konsistensi yang baik dan relevansi yang tinggi terhadap keseluruhan instrumen evaluasi. Hal ini membantu memastikan bahwa tes secara keseluruhan dapat memberikan hasil yang valid dan reliabel.

Tabel 2. Korelasi Skor Butir dengan Skor Total

KORELASI SKOR BUTIR DG SKOR TOTAL
=====

Jumlah Subyek= 30
Butir Soal= 15
Nama berkas: BELUM_ADA_NAMA.ANA

No Butir Baru	No Butir Asli	Korelasi	Signifikansi
1	1	0,532	Signifikan
2	2	0,720	Sangat Signifikan
3	3	0,822	Sangat Signifikan
4	4	0,576	Signifikan
5	5	0,534	Signifikan
6	6	0,728	Sangat Signifikan
7	7	0,323	-
8	8	0,493	Signifikan
9	9	0,441	-
10	10	0,755	Sangat Signifikan
11	11	0,755	Sangat Signifikan
12	12	0,755	Sangat Signifikan
13	13	0,483	Signifikan
14	14	0,744	Sangat Signifikan
15	15	0,659	Sangat Signifikan

Berdasarkan hasil analisis yang dilakukan pada 15 butir soal dengan 30 subjek tes, didapat sebanyak 9 butir soal (nomor 2, 3, 6, 10, 11, 12, 14, dan 15) memiliki korelasi yang sangat signifikan dengan skor total, dengan nilai korelasi di atas 0,65, menunjukkan kontribusi yang sangat baik dalam mengukur kemampuan siswa. Lima butir soal lainnya (nomor 1, 4, 5, 8, dan 13) memiliki korelasi yang signifikan, berkisar antara 0,4–0,6, meskipun masih ada potensi untuk ditingkatkan. Namun, dua butir soal (nomor 7 dan 9) menunjukkan korelasi di bawah 0,4, yang menandakan bahwa butir soal ini kurang efektif dalam berkontribusi terhadap tes dan memerlukan revisi lebih lanjut. Secara keseluruhan, sebagian besar butir soal menunjukkan kualitas yang baik, namun beberapa soal perlu diperbaiki agar tes dapat lebih efektif dalam mengukur kemampuan siswa secara menyeluruh.

Sebuah soal yang memiliki korelasi tinggi dengan skor total menunjukkan bahwa soal tersebut relevan dan efektif dalam mengukur kemampuan yang dimaksudkan oleh keseluruhan tes, yang mengindikasikan bahwa soal tersebut memiliki validitas yang baik. Berdasarkan hasil efektivitas analisis butir soal menggunakan aplikasi anates terhadap korelasi skor butir dengan skor total pada tabel 2 didapat bahwa dari 15 butir soal, 7 soal memiliki kriteria korelasi validitas tinggi dengan angka korelasi antara 0,600–0,800. Sebanyak 6 soal memiliki kriteria korelasi validitas cukup, dengan angka korelasi di antara 0,400–0,600. Satu soal memiliki kriteria korelasi validitas sangat tinggi dengan angka korelasi sebesar 0,822, dan satu soal lainnya memiliki kriteria korelasi validitas rendah dengan angka korelasi 0,323. Hasil ini menunjukkan bahwa sebagian besar butir soal memiliki korelasi yang baik dengan skor total, yang mengindikasikan bahwa soal-soal tersebut relevan dan efektif dalam mengukur kemampuan siswa. Berdasarkan hasil efektivitas analisis butir soal menggunakan aplikasi anates terhadap korelasi skor.

3. Kriteria Daya Beda Soal

Daya pembeda dihitung untuk menentukan kemampuan suatu butir soal dalam membedakan kelompok atas dan kelompok bawah berdasarkan proporsi jawaban benar. Nilai daya pembeda berkisar antara -1 hingga +1. Jika daya pembeda sebesar 0, artinya proporsi peserta kelompok atas dan bawah yang menjawab benar sama besar.

Hanya nilai daya pembeda yang positif yang memiliki arti penting: semakin besar nilainya (mendekati +1), semakin baik kemampuan butir soal tersebut dalam membedakan kelompok atas dan bawah. Hasil analisis daya pembeda dari jawaban peserta didik SMKN 2 Tuban terhadap soal tentang prinsip dan konsep akuntansi dasar dan perbankan dasar dapat dilihat pada tabel 3.

Tabel 3. Daya Pembeda

DAYA PEMBEDA
=====

Jumlah Subyek= 30
Klp atas/bawah(n)= 8
Butir Soal= 15
Nama berkas: BELUM_ADA_NAMA.ANA

No Butir Baru	No Butir Asli	Kel. Atas	Kel. Bawah	Beda	Indeks DP (%)
1	1	8	4	4	50,00
2	2	8	2	6	75,00
3	3	8	1	7	87,50
4	4	8	4	4	50,00
5	5	8	2	6	75,00
6	6	8	1	7	87,50
7	7	8	5	3	37,50
8	8	8	5	3	37,50
9	9	8	4	4	50,00
10	10	8	1	7	87,50
11	11	8	3	5	62,50
12	12	8	1	7	87,50
13	13	8	3	5	62,50
14	14	8	2	6	75,00
15	15	8	4	4	50,00

Berdasarkan hasil analisis, tidak ada butir soal yang berada dalam kategori daya pembeda jelek (0,00–0,20), yang menunjukkan bahwa semua soal memiliki kemampuan diskriminasi yang cukup baik. Sebanyak 4 butir soal (26,67%) berada dalam kategori cukup (0,20–0,40), yaitu soal nomor 3, 6, 7, dan 8 dengan Indeks Daya Pembeda (DP) sebesar 37,50%. Sebanyak 5 butir soal (33,33%) termasuk kategori baik (0,40–0,70), yaitu soal nomor 1, 5, 11, 13, dan 15 dengan DP berkisar antara 50,00% hingga 62,50%. Sementara itu, 6 butir soal (40,00%) masuk dalam kategori sangat baik (0,70–1,00), yaitu soal nomor 2, 4, 9, 10, 12, dan 14 dengan DP sebesar 75,00% hingga 87,50%.

Dengan demikian, dapat disimpulkan bahwa tes ini memiliki kualitas yang cukup baik dalam membedakan kemampuan peserta, karena sebagian besar soal berada dalam kategori baik dan sangat baik. Meskipun terdapat beberapa soal dalam kategori cukup, revisi terhadap soal-soal tersebut dapat dilakukan untuk meningkatkan daya pembeda dan keseluruhan kualitas tes. Hal ini menunjukkan bahwa tes ini dapat digunakan dengan efektif, namun tetap memerlukan penyempurnaan untuk hasil yang lebih optimal.

4. Kriteria Tingkat Kesukaran

Tujuan dari analisis tingkat kesukaran adalah untuk memastikan bahwa soal yang digunakan dapat mencakup beragam tingkat kemampuan siswa. Soal dengan tingkat kesukaran yang moderat dianggap paling ideal, karena dapat menunjukkan perbedaan kemampuan siswa dengan lebih jelas. Sebaliknya, soal yang terlalu mudah atau terlalu sulit kurang efektif dalam menggambarkan kemampuan siswa secara akurat.

Tingkat kesukaran soal biasanya dinyatakan dalam bentuk indeks yang berkisar antara 0,00 hingga 1,00. Semakin besar nilai indeks, semakin mudah tingkat kesukaran soal tersebut. Indeks kesukaran digunakan untuk menggambarkan seberapa banyak peserta didik yang dapat menjawab soal dengan benar.

Tabel 4. Tingkat Kesukaran

TINGKAT KESUKARAN

=====

Jumlah Subyek= 30
Butir Soal= 15
Nama berkas: BELUM_ADA_NAMA.ANA

No Butir Baru	No Butir Asli	Jml Betul	Tkt. Kesukaran(%)	Tafsiran
1	1	25	83,33	Mudah
2	2	19	63,33	Sedang
3	3	18	60,00	Sedang
4	4	25	83,33	Mudah
5	5	19	63,33	Sedang
6	6	17	56,67	Sedang
7	7	26	86,67	Sangat Mudah
8	8	24	80,00	Mudah
9	9	21	70,00	Sedang
10	10	22	73,33	Mudah
11	11	22	73,33	Mudah
12	12	22	73,33	Mudah
13	13	19	63,33	Sedang
14	14	21	70,00	Sedang
15	15	26	86,67	Sangat Mudah

Tingkat kesukaran soal dapat dibagi menjadi tiga kategori berdasarkan persentase jawaban benar peserta. Soal Mudah termasuk butir soal nomor 1, 4, 7, 8, 10, 11, 12, dan 15, dengan persentase jawaban benar antara 73,33% hingga 86,67%. Soal-soal ini dapat dianggap relatif mudah karena sebagian besar peserta mampu menjawabnya dengan benar. Soal Sedang terdiri dari butir soal nomor 2, 3, 5, 6, 9, 13, dan 14, dengan persentase jawaban benar sekitar 56,67% hingga 70%. Soal-soal ini menunjukkan variasi dalam kemampuan peserta, dengan beberapa peserta menjawab dengan benar dan lainnya kesulitan. Sementara itu, Soal Sangat Mudah hanya terdapat pada butir soal nomor 7 dan 15, dengan persentase jawaban benar mencapai 86,67%, menunjukkan bahwa sebagian besar peserta dapat dengan mudah menjawab soal-soal ini.

5. Rangkuman Hasil Analisis Butir Soal

Soal tes merupakan salah satu alat evaluasi yang penting dalam pendidikan, digunakan untuk mengukur pencapaian pembelajaran siswa pada berbagai tingkat kognitif (Sulistiawan, 2016). Tes berfungsi sebagai sarana untuk mengetahui sejauh mana siswa memahami materi yang telah diajarkan, mengidentifikasi kelemahan dalam proses belajar-mengajar, serta sebagai dasar untuk perbaikan pembelajaran.

Hasil analisis soal tes yang baik memiliki validitas tinggi, reliabilitas yang handal, tingkat kesukaran soal tes pada kondisi sedang, dan daya beda soal tes dengan tingkatan yang baik tiap soal tes (Susanto, 2013). Soal yang ideal memiliki validitas tinggi, artinya soal-soal tersebut benar-benar mengukur kompetensi atau materi yang dituju dengan akurat. Selain itu, tes juga harus memiliki reliabilitas yang baik, yakni kemampuan untuk menghasilkan hasil yang konsisten dan stabil ketika diterapkan pada kondisi atau kelompok yang berbeda. Tingkat kesukaran soal sebaiknya berada pada kategori sedang (indeks kesukaran antara 0,30–0,70), sehingga memungkinkan perbedaan kemampuan siswa terlihat secara jelas. Terakhir, soal harus memiliki daya beda yang baik, yaitu kemampuan untuk membedakan siswa yang memiliki kemampuan tinggi dan rendah.

Tabel 5. Rangkuman Hasil Analisis Butir Soal

NO.	Kriteria Validitas	Kriteria Daya Beda Soal	Kriteria Tingkat Kesukaran	Koefisien Reliabilitas	Keputusan
1	Cukup	Baik	Mudah	0,92	Revisi
2	Tinggi	Baik	Sedang	0,92	Digunakan
3	Sangat Tinggi	Sangat Baik	Sedang	0,92	Digunakan
4	Cukup	Baik	Mudah	0,92	Revisi
5	Cukup	Sangat Baik	Sedang	0,92	Digunakan
6	Tinggi	Sangat Baik	Sedang	0,92	Digunakan
7	Rendah	Cukup	Sangat Mudah	0,92	Revisi
8	Cukup	Cukup	Mudah	0,92	Revisi
9	Cukup	Baik	Sedang	0,92	Digunakan
10	Tinggi	Sangat Baik	Mudah	0,92	Digunakan
11	Tinggi	Baik	Mudah	0,92	Digunakan

12	Tinggi	Sangat Baik	Mudah	0,92	Digunakan
13	Cukup	Baik	Sedang	0,92	Digunakan
14	Tinggi	Sangat Baik	Sedang	0,92	Digunakan
15	Tinggi	Baik	Sangat Mudah	0,92	Digunakan

Analisis terhadap kriteria validitas, daya beda soal, tingkat kesukaran, dan koefisien reliabilitas pada data yang diberikan menunjukkan bahwa sebagian besar soal memiliki kualitas yang baik. Dari 15 soal yang dievaluasi, terdapat 11 soal yang dinyatakan "Digunakan" dan 4 soal lainnya perlu direvisi.

Dari segi validitas, 5 soal memiliki kriteria "cukup," yang menunjukkan bahwa meskipun validitasnya tidak optimal, masih ada potensi untuk perbaikan. Sebaliknya, 10 soal lainnya memiliki validitas yang "tinggi" atau "sangat tinggi," yang menandakan bahwa soal-soal tersebut dapat diandalkan untuk mengukur kemampuan peserta.

Dalam hal daya beda soal, sebagian besar soal dikategorikan sebagai "Baik" atau "Sangat Baik," yang menunjukkan bahwa soal-soal tersebut mampu membedakan antara peserta yang memiliki kemampuan tinggi dan rendah. Namun, terdapat 2 soal yang memiliki daya beda "Cukup," yang mungkin memerlukan perhatian lebih.

Tingkat kesukaran soal bervariasi, dengan beberapa soal dikategorikan sebagai "Mudah" dan "Sedang." Soal-soal yang terlalu mudah atau terlalu sulit mungkin tidak efektif dalam mengevaluasi kemampuan peserta secara akurat. Dalam hal ini, 5 soal yang dinyatakan "Mudah" perlu direvisi untuk meningkatkan tingkat kesukaran yang sesuai.

Koefisien reliabilitas untuk semua soal adalah 0,92, yang menunjukkan bahwa soal-soal tersebut memiliki tingkat konsistensi yang tinggi. Hal ini berarti bahwa hasil yang diperoleh dari soal-soal ini dapat diandalkan. Secara keseluruhan, meskipun sebagian besar soal dapat digunakan, ada beberapa yang perlu direvisi untuk meningkatkan kualitas dan efektivitas penilaian.

KESIMPULAN

Evaluasi dalam pembelajaran merupakan tolok ukur yang digunakan untuk mengetahui kemampuan belajar peserta didik. Alat evaluasi dapat berupa tes dengan format pilihan ganda, menjodohkan, atau isian singkat. Dalam penelitian ini, digunakan soal berbentuk pilihan ganda yang diujikan kepada peserta didik. Untuk menghasilkan soal yang berkualitas dan memenuhi kriteria tertentu, diperlukan analisis tes. Anates adalah software yang dirancang untuk membantu guru atau pendidik dalam menganalisis kualitas soal, baik berupa pilihan ganda maupun soal uraian. Aplikasi ini memudahkan guru untuk memasukkan data nilai peserta didik dan kunci jawaban ke dalam sistem. Selanjutnya, Anates akan mengolah data tersebut dan menghasilkan informasi berupa reliabilitas soal, identifikasi kelompok atas (unggul) dan kelompok bawah (rendah), tingkat kesukaran butir soal, kualitas pengecoh, daya pembeda, serta korelasi antara bobot skor dengan skor total.

Berdasarkan data, sebagian besar butir soal memiliki korelasi signifikan dengan skor total, seperti soal no 3 (0,822) dan no 10 (0,755), yang menunjukkan validitas baik, meskipun soal no 7 (0,323) memiliki korelasi rendah. Daya pembeda soal berkisar antara 37,50% (no 7 dan 8) hingga 87,50% (no 3, 6, 10, dan 12), di mana soal dengan daya pembeda tinggi efektif membedakan kemampuan peserta, sementara soal dengan daya pembeda rendah perlu evaluasi. Tingkat kesukaran soal bervariasi dari sangat mudah (no 7 dan 15, indeks 0,8667) hingga sedang (no 3, indeks 0,60), menunjukkan distribusi baik untuk mengukur berbagai kemampuan. Dengan reliabilitas tes sebesar 0,92, tes ini memiliki konsistensi tinggi dan dapat dipercaya.

DAFTAR PUSTAKA

Alpusari, M. (2014). Analisis Butir Soal, Program Komputer Anates Versi 4.0 For Windows. *Jurnal Primary Program Studi Pendidikan Guru Sekolah Dasar Universitas Riau*, 3(2), 107.

- <https://pdfs.semanticscholar.org/0cb3/b3e65e93596f7007710b2dd229f1ea2a5205.pdf>
- Arif, M. (2015). Penerapan Aplikasi Anates Bentuk Soal Pilihan Ganda. *Edutic - Scientific Journal of Informatics Education*, 1(1), 1–9. <https://doi.org/10.21107/edutic.v1i1.398>
- Arikunto, S. (2013). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta : Bumi Aksara.
- Cahyaningrum, I. Y., Fuady, A., & Sunismi. (2023). Analisis Butir Soal Sumatif Akhir Semester Ganjil Mata Pelajaran Matematika Kelas VII dengan Berbantuan Aplikasi Software Anates. *Mathema Journal*, 5(2), 67–81.
- Fibrina Aquatika, Rufi'I, & Moch. Hendy Bayu Pratama. (2022). Penggunaan Aplikasi Anates pada Analisis Butir Soal Penilaian Tengah Semester Bahasa Indonesia Kelas X MIPA 8 SMA Muhammadiyah 2 Surabaya. *IKIP Widya Darma Surabaya*, 4, 1–7. <http://jurnal.ikipwidyardarmasurabaya.ac.id/index.php/metalanguage/login>
- Fitriani, N. (2021). Analisis Tingkat Kesukaran, Daya Pembeda, Dan Efektivitas Pengecoh Soal Pelatihan Kewaspadaan Kegawatdaruratan Maternal Dan Neonatal. *Paedagogia : Jurnal Kajian, Penelitian Dan Pengembangan Kependidikan*, 12(2), 199. <https://doi.org/10.31764/paedagogia.v12i2.4956>
- Helpiana, N. (2020). Analisis Butir Soal Ujian Akhir Semester Genap Tahun Ajaran 2018/2019 Mata Pelajaran PAI Kelas XI SMAN 2 Gunung Meriah Kabupaten Aceh Singkil.
- Kaka, L., Bano, V. O., & Njoeroemana, Y. (2024). Efektivitas Analisis Butir Soal Pilihan Ganda Menggunakan Aplikasi Anates Di Smpn 2 Kandatng. *Jurnal Inovasi Penelitian*, 1441(1450), 4–9. <https://stp-mataram.e-journal.id/JIP/article/view/3124/2473>
- L, I. (2019). EVALUASI DALAM PROSES PEMBELAJARAN Idrus L 1. *Evaluasi Dalam Proses Pembelajaran*, 9(2), 344.
- M Ilyas Ismail, D. (2020). Asesmen dan Evaluasi Pembelajaran. <https://books.google.co.id/books?id=aUvODwAAQBAJ&lpg=PP1&dq=pengertian evaluasi&lr&pg=PP1#v=onepage&q=pengertian evaluasi&f=false>
- M. B. (2017). Evaluasi Belajar Peserta Didik (Siswa). *Idaarah: Jurnal Manajemen Pendidikan*, 1(2), 257–267. <https://doi.org/10.24252/idaarah.v1i2.4269>
- Mochammad Noor Akhmadi. (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>
- Nurwinda, R. (2013). Analisis Penggunaan Alat Evaluasi Pada Mata Pelajaran Produktif Akuntansi Kelas X Akuntansi Di Smk Negeri 2 Buduran Sidoarjo. *Jurnal Pendidikan Akuntansi & Keuangan*, 1. <https://doi.org/10.17509/jpak.v2i1.15455>
- Pamungkasih, R. S. N., & Nawawi, E. (2021). Analisis Kualitas Butir Soal Ujian Akhir Semester Ganjil pada Mata Pelajaran Kimia Kelas X di SMA Negeri 8 Palembang Tahun Ajaran 2020/2021. *Prosiding Seminar Nasional Pendidikan IPA Tahun 2021*, 1.
- Putri, D., & Retnosari, D. S. (2024). Analisis Kualitas Soal Penilaian Akhir Semester Mata Pelajaran Dasar Program Keahlian Akuntansi dan Keuangan Lembaga di SMK Ketintang Surabaya. 7(April), 8–17.
- Sudijono, A. (2011). *Pengantar Evaluasi Pendidikan*. Jakarta : PT Raja Grafindo Persada. <https://id.scribd.com/document/413832352/Prof-Drs-Anas-Sudijono-Pengantar-Evaluasi-Pendidikan-intro-pdf>