
Analisis Butir Soal HOTS untuk Menguji Pemahaman Siswa SMK Jurusan Akuntansi Menggunakan *Software Anates*

Dinda Meilin Syafiriya
Universitas Negeri Surabaya

Putri Ayu Rahmawati
Universitas Negeri Surabaya

Zuhrotul Qolbi
Universitas Negeri Surabaya

Luqman Hakim
Universitas Negeri Surabaya

Vivi Pratiwi
Universitas Negeri Surabaya

Alamat :

Program Studi Pendidikan Akuntansi, Fakultas Ekonomika dan Bisnis
Universitas Negeri Surabaya
Jl. Ketintang, Ketintang, Kec. Gayungan, Kota Surabaya, Jawa Timur 60231
Korespondensi penulis: meilindinda@gmail.com

Abstract. *This question trial research aims to determine the understanding of vocational students majoring in Financial and Institutional Accounting in the Accounting Practicum elements of Service, Trading and Manufacturing Companies conducted in November 2024. This research was conducted online by distributing g-forms to research subjects as many as 30 students. The data analysis technique used the Anates version 4.0.9 program to determine the reliability, difficulty level, differentiating power, quality of checkers, and item analysis recap. The results showed that, the reliability value of the test questions had sufficient criteria with a reliability value of 0.65. For the level of difficulty of the items, it is known that 12 questions with the level of difficulty "Very easy" and 3 questions with the level of difficulty "easy". In the differentiating power of the items, there are 8 items classified as "Poor", 4 items classified as "Fair", 2 items classified as "Good", and 1 item classified as "Very Good". Therefore it can be concluded that, there are 4 items that can be used but still have to be improved and there are 8 items that must be improved. So that only 3 items out of 15 items can be used for student test materials.*

Keywords: *Research, Trial, HOTS Problem Items, Accounting Practicum for Service, Trading, and Manufacturing Companies.*

Abstrak. Penelitian uji coba soal ini bertujuan untuk mengetahui pemahaman siswa SMK jurusan Akuntansi Keuangan dan Lembaga pada elemen Praktikum Akuntansi Perusahaan Jasa, Dagang, dan Manufaktur yang dilakukan pada bulan November 2024. Penelitian ini dilakukan secara online dengan menyebarkan g-form kepada subjek penelitian sebanyak 30 siswa. Analisis data dilakukan dengan program Anates versi 4.0.9 untuk mengukur reliabilitas, kesulitan, daya beda, kualitas pengecoh, serta rekapitulasi analisis item soal. Hasil penelitian menunjukkan bahwa nilai reliabilitas dari soal tes berkriteria cukup dengan nilai reliabilitas sebesar 0,65 . Untuk Tingkat kesulitan butir soal diketahui 12 soal dengan tingkat kesulitan “Sangat mudah” dan 3 soal tingkat kesulitan “mudah”. Dalam daya pembeda butir soal, terdapat 8 butir soal yang termasuk dalam kategori “Buruk”, 4 butir soal yang tergolong “Cukup”, 2 butir soal yang masuk kategori “Baik”, dan 1 butir soal yang tergolong “Sangat Baik”. Oleh karena itu dapat disimpulkan bahwa, terdapat 4 butir soal yang dapat digunakan namun masih harus diperbaiki dan terdapat 8 butir soal yang harus diperbaiki. Sehingga hanya tersisa 3 butir soal dari 15 soal yang dapat digunakan untuk bahan uji coba peserta didik.

Kata kunci: Penelitian, Uji Coba, Butir Soal HOTS, Praktikum Akuntansi Perusahaan Jasa, Dagang, dan Manufaktur.

LATAR BELAKANG

Pendidikan sendiri adalah rangkaian sesuatu yang dilakukan individu, kelompok, lembaga, dan lain-lain dalam rangka mencapai tahapan pendidikan. Proses itu sendiri melibatkan sistem pendidikan, yang memiliki sejumlah elemen yang saling berhubungan, dan bahwa tidak dapat berfungsi tanpa keberadaan satu sama lain. Peserta didik, guru, kurikulum, serta satu sama lain memiliki hubungan yang saling berhubungan dan bekerja sama (Leonardus K. dkk., 2024). Peran yang vital dari semua faktor tersebut ialah evaluasi sebagai prosedur membuktikan pencapaian pembelajaran hasil siswa. Oleh sebab itu, mutlak dibutuhkan bagi seorang guru. Salah satu elemen krusial di kegiatan belajar mengajar adalah evaluasi, biasanya dilakukan dengan cara ujian atau tes.

Tes adalah alat yang dirancang untuk mengukur suatu kemampuan belajar siswa. Ini terdiri dari serangkaian pertanyaan atau tugas (Muchlizani. N. dkk., 2023). Jadi dengan adanya tes, maka kita dapat mengukur kemampuan dari peserta didik. Dalam metode analisis butir soal, pertimbangan utama adalah validitas dan reliabilitas soal. Meskipun demikian, ketika mengevaluasi jenis soal objektif seperti pilihan ganda, analisis juga memperhitungkan tingkat kesulitan, daya pembeda, dan efektivitas

pengalih perhatian. (Ramadhan. W. dkk., 2023). Maka agar dapat menghasilkan ujian dengan kualitas yang baik, maka analisis butir soal sangat diperlukan.

Analisis butir soal untuk menguji mutu setiap butir soal dan seperangkat soal dari aspek kuantitas dan kualitas. Analisis butir soal bisa ditempuh dengan dua cara, yaitu dengan telaah butir soal dan dengan analisis empiris (Elviana, 2020). Baik analisis terhadap butir soal maupun peninjauan berdasarkan data, bermanfaat untuk memperoleh informasi mengenai karakteristik soal. Beberapa temuan ini bisa dimanfaatkan untuk menilai kualitas soal serta kualitas pembelajaran siswa berdasarkan hasil analisis setiap butir soal. Seorang guru harus memiliki kemampuan untuk membuat dan menyusun alat penilaian dalam bentuk pertanyaan yang ditujukan untuk mengevaluasi kemampuan berpikir tingkat tinggi siswa, yang umumnya disebut sebagai Keterampilan Berpikir Tingkat Tinggi (HOTS) (Gradini. E. dkk., 2019). Maka dengan guru yang bisa menganalisis butir soal yang dibuatnya, butir soal tersebut pasti akan bisa mengukur kemampuan berpikir peserta didik.

HOTS adalah keterampilan berpikir tingkat tinggi yang mengacu pada ranah C4 (analisis), C5 (evaluasi), dan C6 (kreasi) (Saraswati. P. M. S. dkk., 2020). Soal-soal HOTS yang telah dirancang secara efektif mengharuskan siswa untuk melakukan pemikiran kritis, logis, reflektif, metakognitif, dan kreatif. Saat mengembangkan pertanyaan HOTS, beberapa faktor harus diperhitungkan, termasuk pembuatan kisi-kisi pertanyaan, pemilihan stimulus yang menarik dan kontekstual, dan penetapan kriteria penilaian (Widhiyani. I. A. N. T. dkk., 2019). Setelah semua sudah dipertimbangkan maka akan dikatakan bisa digunakan untuk mengukur kemampuan peserta didik melalui HOTS. Anates dapat menganalisis soal tes dengan menghitung skor, realibilitas tes, tingkat kesukaran, daya pembeda, kelompokkan subjek ke dalam kelompok unggul dan asor, kualitas pengecoh (disktaktor), dan menggunakan sistem operasi Windows (Amelia. N. dkk., 2021). Maka anates sangat membantu guru dalam menganalisis butir soal untuk mengetahui seberapa berpikir kritisnya peserta didik dan mengukur kemampuan dalam penguasaan mata pelajaran tersebut.

KAJIAN TEORITIS

Pendidikan adalah salah satu elemen yang begitu penting dalam proses pembangunan di suatu bangsa, di mana evaluasi juga memiliki peran yang sangat

krusial untuk mengukur prestasi belajar siswa. Ujian atau tes adalah salah satu bentuk evaluasi yang biasa dipergunakan dalam suatu pendidikan (Yusuf. W. F. dkk., 2024). Maka dari itu, pendidikan menjadi kunci pertama dalam pegangan untuk si masa yang akan mendatang. Dengan adanya berbagai soal untuk mengetes pendidikan di dalam peserta didik, maka untuk mengukur kemampuan peserta didik perlu adanya menganalisis butir soal. Menurut Lestari, Wardana, & Rahmawati (2019) Analisis butir soal adalah tugas yang harus segera dilakukan oleh guru untuk meningkatkan kualitas soal. Jadi dengan menganalisis butir soal, kita dapat tahu seberapa sulit dan mudah dalam suatu soal.

Menganalisis butir soal sangat mudah jika menggunakan *software* Anates. Anates adalah salah satu dari banyak program yang dimanfaatkan untuk menilai kualitas setiap butir soal dengan menggunakan pendekatan klasik atau analisis. Program ini mengevaluasi validitas, reliabilitas, tingkat kesukaran, daya pembeda, serta keefektifan pengecoh soal dengan cepat (Fiska. J. M. dkk., 2021). Anates membuat guru menjadi mudah dalam menganalisis soal yang dibuat itu termasuk soal yang tingkat mudah atau sulit, bahkan bisa diambil kesimpulan soal itu akan di perbaiki ataupun bisa digunakan.

Di dalam pertanyaan untuk menganalisis butir soal, maka butir soal tersebut menggunakan butir soal yang HOTS. Menurut Saraswati. P. M. S. dkk. (2020) HOTS adalah sebagai kemampuan untuk memecahkan masalah dengan menggunakan pemikiran kritis dan kreatif. Maka untuk menemukan solusi atau ide baru, seseorang perlu memiliki keterampilan berpikir yang canggih untuk melakukan analisis, menghubungkan, menguraikan, dan memaknai masalah. Dalam revisi Taksonomi Bloom , HOTS berada pada tingkat menganalisis, mengevaluasi, dan mencipta. Menurut Pramarth, I. N. B., dkk. (2023) HOTS juga merupakan bagian dari ranah kognitif yang ada. Misalnya, untuk dapat menyelesaikan suatu masalah, memikirkan solusi alternatif, menggunakan metode penyelesaian masalah, dan mengevaluasi strategi dan solusi yang digunakan.

Di *software* Anates, juga dapat menganalisis reliabilitas tinggi menunjukkan bahwa soal tersebut memberikan hasil yang konsisten jika digunakan dalam berbagai tes. Selain itu, bisa menganalisis daya pembeda dalam suatu soal dengan soal lainnya. Daya pembeda merujuk pada seberapa efektif sebuah soal dalam membedakan antara siswa yang sangat terampil dan kurang terampil. Dengan kata lain, daya pembeda ini

mengukur seberapa baik soal tersebut mampu dalam menampilkan perbedaan penguasaan materi di antara siswa dengan kemampuan serupa dibandingkan dengan siswa yang memiliki kemampuan lebih rendah. Di dalam software Anates juga dapat menganalisis kualitas pengecoh di setiap butir soal. Menurut Ratih, P. D., dkk. (2024) Kualitas pengecoh adalah proses evaluasi yang bertujuan untuk menganalisis setiap pilihan jawaban yang tersedia. Jadi analisis kualitas pengecoh mengidentifikasi pilihan yang efektif dan layak dipertahankan, dan pilihan yang tidak tepat yang memerlukan perbaikan atau penggantian.

Setelah semua sudah di analisis, maka ditentukannya kelompok unggul dan asor. Kelompok unggul dan asor adalah istilah yang digunakan dalam analisis butir soal agar bisa membedakan siswa berdasarkan kemampuan mereka. Kelompok siswa dengan skor tertinggi berada di atas, dan kelompok siswa dengan skor rendah berada di bawah. (Ambarwati & Ismiyati, 2021). Jadi metode pengelompokan analisis butir soal membagi peserta tes menjadi dua kelompok berdasarkan skor total mereka.

Menurut peneliti yang dilakukan Yusuf W F (2024) penelitian tentang analisis butir soal ini sangat relevan dan bermanfaat untuk diterapkan dalam dunia pendidikan. Analisis butir soal dengan menggunakan Anates untuk menentukan kelayakan butir soal, terutama untuk jenis soal pilihan ganda. Ini mencakup validitas, reliabilitas, tingkat kesukaran, daya pembeda dan kualitas pengecoh. Serta sama halnya menurut Fiska, Joelan M., Yunin H., Nur Qomaria., dan Wiwin Puspita H. (2021) Anates juga salah satu jenis software yang dapat menganalisis kualitas setiap butir soal dengan menggunakan pendekatan klasik atau ditinjau. Ini melakukannya dengan mempertimbangkan validitas, reliabilitas, tingkat kesukaran, daya beda, dan efektivitas pengecoh soal dengan cepat. Anates juga memiliki kemampuan untuk menganalisis bagian soal dalam bentuk pilihan ganda dan uraian. Sehingga lebih mudah bagi pengguna, Anates sendiri juga berbahasa Indonesia.

METODE PENELITIAN

Penelitian ini menggunakan penelitian kuantitatif dengan metode deskriptif. Melalui penelitian deskriptif kuantitatif, kita dapat menyajikan data dalam bentuk angka untuk menggambarkan secara jelas kondisi, situasi, atau kejadian yang sedang diteliti. Penelitian ini merupakan Penelitian yang digunakan sebagai mengukur dan menguji keefektifan butir soal HOTS pada siswa SMK jurusan Akuntansi Keuangan dan

Lembaga terhadap elemen Praktikum Akuntansi Perusahaan Jasa, Dagang, dan Manufaktur di Kurikulum Merdeka. Populasi di Penelitian ini, terdiri dari 30 siswa SMK jurusan Akuntansi Keuangan dan Lembaga. Metode pengambilan sampel yang ada di penelitian ini yaitu dengan menggunakan teknik tes.

Metode pengumpulan data yang digunakan di penelitian ini meliputi menggunakan g-form yang disebar ke Siswa sebagai sampel Penelitian. Metode yang digunakan pada analisis data adalah memasukkan data ke aplikasi Anates kemudian dianalisis untuk menghasilkan skor data yang berbobot, reabilitas tes, kelompok unggul dan biasa, daya pembeda, tingkat kesulitan, korelasi skor butir dengan skor total, dan kualitas pengecoh. Penelitian ini merupakan Penelitian kuantitatif dengan menggunakan *software Anates for Windows*. Metode Penelitian ini antara lain:

1. Tahap Persiapan

- a) Menetapkan elemen soal yang akan dibuat.
- b) Membuat soal HOTS yang akan diteliti di G-Form.
- c) Menginstal *software Anates* yang digunakan untuk menganalisis.

2. Tahap Pelaksanaan

- a) Menyebar soal yang akan diujikan ke siswa.
- b) Memeriksa hasil jawaban siswa.
- c) Memasukkan hasil jawaban siswa ke *software Anates*.

3. Analisis dan Kriteria Keberhasilan

a) Skor yang Dibobot

Skor yang dibobot akan menunjukkan hasil jawaban dari peserta didik yang menjawab benar dan salah, skor asli, dan skor bobot pada masing-masing peserta tes.

b) Analisis Reabilitas

Uji reabilitas digunakan untuk memastikan apakah suatu alat ukur, misalnya tes atau kuesioner, akan memunculkan hasil yang sama atau konsisten jika digunakan berulang dalam kondisi yang sama. Uji reabilitas bertujuan untuk menilai stabilitas alat ukur, atau memastikan bahwa alat ukur yang digunakan tersebut menghasilkan hasil yang dapat diandalkan dan konsisten

pada berbagai pengukuran (Fietri. W. A. dkk., 2021). Kriteria reabilitas dapat dilihat pada tabel 1.

Tabel 1. Kriteria Realibilitas

Batasan	Interpretasi
0,00 – 0,20	Sangat Rendah
0,20 – 0,40	Rendah
0,40 – 0,70	Cukup
0,70 – 0,90	Tinggi
0,90 – 1,00	Sangat Tinggi

c) **Tingkat Kesukaran**

Pada tingkat kesukaran soal menunjukkan seberapa banyak peserta tes yang berhasil atau bisa menjawab soal dengan tepat dan benar. Semakin tinggi persentase siswa dalam menjawab soal dengan tepat, maka semakin mudah soal tersebut, sebaliknya jika semakin rendah persentase siswa dalam menjawab soal dengan tepat maka semakin sulit soal itu (Saputri, H. A. S., & Larasati, N. J., 2023). Sebuah tes yang baik memiliki soal-soal yang bervariasi dalam tingkat kesulitan yang seimbang, tidak terlalu banyak soal yang sangat mudah atau sangat sulit. Setiap soal harus memiliki tingkat kesulitan 25%, sedang 25%, dan sukar 25% saat menyusun soal tes (Ashari. S. D. dkk., 2021). Kriteria tingkat kesukaran dapat kita lihat pada tabel 2.

Tabel 2. Kriteria Tingkat Kesukaran

Batasan	Interpretasi
0,00 – 0,30	Sukar
0,31 – 0,70	Sedang
0,71 – 1,00	Mudah

d) **Daya Pembeda**

Yang di maksud daya pembeda suatu soal yaitu kemampuan dari butir soal tersebut untuk memisahkan peserta didik yang memiliki tingkat kemampuan tinggi (kelompok atas) dan memiliki kemampuan rendah (kelompok bawah) dalam memahami materi yang diujikan (Pradita. E. dkk., 2023). Tingkat kemampuan soal ini diukur dengan angka yang disebut daya pembeda atau indeks diskriminasi. Kriteria daya pembeda dapat kita lihat pada tabel 3.

Tabel 3. Kriteria Daya Pembeda

Batasan	Interpretasi
0,00 – 0,20	Jelek
0,20 – 0,40	Cukup
0,40 – 0,70	Baik
0,70 – 1,00	Sangat Baik

e) Kelompok Unggul dan Asor

Menunjukkan kelompok peserta yang menjadi bagian kelompok unggul (Kelompok Atas) dan bagian kelompok asor (Kelompok Bawah).

f) Korelasi Skor Butir dengan Skor Total

Menunjukkan seberapa jauh jawaban pada butir soal terkait dengan pencapaian peserta secara keseluruhan. Kriteria korelasi skor butir dengan skor total dapat kita dilihat pada tabel 4.

Tabel 4. Kriteria Korelasi Validitas

Batasan	Interpretasi
0,800 – 1,00	Sangat Tinggi
0,600 – 0,800	Tinggi
0,400 – 0,600	Cukup
0,200 – 0,400	Rendah
0,000 – 0,200	Sangat Rendah

g) Kualitas Pengecoh

Menunjukkan alternatif (Option atau Pilihan) jawaban yang baik digunakan kembali dan alternatif jawaban yang buruk agar diubah atau diganti. Kriteria kualitas pengecoh dapat kita dilihat pada tabel 5.

Tabel 5. Kriteria Kualitas Pengecoh

Keterangan	Interpretasi
++	Sangat Baik
+	Baik
-	Kurang Baik
--	Buruk

Sangat Buruk

h) Rekap Analisis Butir Soal

Menunjukkan rekap hasil analisis butir soal yang pada akhirnya akan diambil keputusan, yaitu soal diperbaiki, soal diperbaiki dan digunakan, dan soal digunakan.

HASIL DAN PEMBAHASAN

Analisis butir soal yaitu proses yang dilakukan oleh pendidik untuk memeriksa dan memperbaiki kualitas soal sebelum diberikan kepada siswa. Tujuannya adalah untuk memastikan bahwa soal yang digunakan berkualitas tinggi dan dapat dengan akurat mengukur kemampuan siswa. Hasil penelitian ini menunjukkan bahwa kualitas suatu tes dapat dinilai berdasarkan :

1. Skor Data Dibobot

Skor data dibobot merupakan hasil pengolahan nilai mentah yang telah dikonversi. Skor data dibobot menampilkan hasil dari respons siswa yang memberikan jawaban benar dan salah, skor asli, dan skor bobot pada setiap siswa yang mengikuti tes (Ratih, P. D., dkk., 2024).

Gambar 1. Skor Data Dibobot

No Urt	No Subyek	Kode/Nama	Benar	Salah	Kosong	Skor Asli	Skor Bobot
1	2	CYNTIA...	15	0	0	15	15
2	4	AQILA ...	15	0	0	15	15
3	5	JUNI T...	15	0	0	15	15
4	7	ANDINI...	15	0	0	15	15
5	11	IRDA S...	15	0	0	15	15
6	14	KHOIRU...	15	0	0	15	15
7	29	RAFIF ...	15	0	0	15	15
8	1	KHOLIS...	14	1	0	14	14
9	3	DIVA Y...	14	1	0	14	14
10	6	AIS BA...	14	1	0	14	14
11	9	ELOK M...	14	1	0	14	14
12	10	FAHRUL...	14	1	0	14	14
13	12	LIDYA ...	14	1	0	14	14
14	16	FAULA ...	14	1	0	14	14
15	21	RINDI ...	14	1	0	14	14
16	23	SARAH ...	14	1	0	14	14
17	25	FANNY ...	14	1	0	14	14
18	26	DEANDR...	14	1	0	14	14
19	8	IIN NU...	13	2	0	13	13
20	13	CALLYS...	13	2	0	13	13
21	15	FEBRIL...	13	2	0	13	13
22	17	CHINDY...	13	2	0	13	13
23	22	LATIFA...	13	2	0	13	13
24	27	NENSYA...	13	2	0	13	13
25	18	SASKIA...	12	3	0	12	12
26	19	RAHMA ...	12	3	0	12	12
27	20	THELAN...	12	3	0	12	12
28	24	ERLIND...	12	3	0	12	12
29	30	ZEBINA...	10	5	0	10	10
30	28	DZUL'I...	6	9	0	6	6

Tabel 6. Hasil Skor Data Dibobot

Skor	Jumlah	Persentase%
15	7	23,3%

14	11	36,6%
13	6	20%
12	4	13,3%
10	1	3,3%
6	1	3,3%

Tabel 6 menunjukkan skor yang diperoleh siswa yang digunakan untuk acuan yang menilai kemampuan mereka dalam berpikir tingkat tinggi. Kemampuan ini mencakup kemampuan dalam menyelesaikan masalah dan membuat keputusan. Hasil skor 15 sebanyak 7 (23,3%) siswa, skor 14 sebanyak 11 (36,6%) siswa, skor 13 sebanyak 6 (20%) siswa, skor 12 sebanyak 4 (13,3%) siswa, skor 10 dan skor 6 masing-masing sebanyak 1 (3,3%) Siswa.

2. Analisis Realibilitas

Hasil uji reabilitas butir soal berupa pilihan ganda menunjukkan sebesar 0,65. Analisis reliabilitas uji coba butir soal menggunakan *software* Anates menunjukkan bahwa nilai reliabilitasnya berada di antara 0,40 – 0,70. Rentang nilai ini mengindikasikan bahwa pertanyaan-pertanyaan tersebut memiliki tingkat reabilitas yang "cukup". Namun, untuk dianggap sebagai soal yang baik, nilai reliabilitasnya seharusnya lebih tinggi, yaitu di atas 0,70. Hasil reabilitas ditunjukkan pada tabel 7 sebagai berikut :

Tabel 7. Hasil Reabilitas

Rata-Rata	13,37
Simpang Baku	1,83
Korelasi XY	0,48
Reabilitas	0,65

Tabel 7 menunjukkan koefisien reliabilitas soal uji coba adalah 0,65. Nilai-nilai ini menandakan bahwa pertanyaan-pertanyaan tersebut memiliki tingkat konsistensi “cukup” dalam mengukur kemampuan siswa yang masih perlu ditingkatkan. Dari hasil analisis, kita bisa simpulkan bahwa soal ini kurang bisa diandalkan. Artinya, soal-soal ini belum benar-benar bisa mengukur kemampuan siswa secara akurat. Jadi, nilai yang didapat siswa belum tentu mencerminkan kemampuan sebenarnya.

Kualitas reliabilitas suatu tes sangat dipengaruhi oleh berbagai faktor. Mulai dari proses pelaksanaan tes (misalnya, kejelasan petunjuk, kesiapan peserta, dan

kondisi tempat tes), jumlah peserta yang mengikuti tes, hingga tingkat kesulitan soal. Jika pelaksanaan tes kurang optimal, variasi jawaban peserta tinggi, atau soal terlalu sulit, maka reliabilitas tes cenderung menurun. Peserta dapat menjadi frustrasi dengan soal yang terlalu sulit dan menebak jawaban secara acak, sehingga hasil tes menjadi kurang akurat.

3. Tingkat Kesukaran Butir Soal

Berdasarkan analisis dengan menggunakan *software* Anates, beberapa soal dapat dikelompokkan menjadi tiga kategori tingkat kesukaran yaitu mudah dan sangat mudah. Hasil analisis tingkat kesukaran dapat dilihat sebagai berikut :

Gambar 2. Tingkat Kesukaran Butir Soal

```

TINGKAT KESUKARAN
=====

Jumlah Subyek= 30
Butir Soal= 15
Nama berkas: C:\USERS\ASUS\ONEDRIVE\KULIAH\MATA KULIAH SEMESTER 3\EVALUASI E

  No Butir Baru  No Butir Asli  Jml Betul  Tkt. Kesukaran(%)  Tafsiran
    1             1           28           93,33  Sangat Mudah
    2             2           24           80,00  Mudah
    3             3           27           90,00  Sangat Mudah
    4             4           22           73,33  Mudah
    5             5           26           86,67  Sangat Mudah
    6             6           27           90,00  Sangat Mudah
    7             7           30          100,00  Sangat Mudah
    8             8           26           86,67  Sangat Mudah
    9             9           27           90,00  Sangat Mudah
   10            10           27           90,00  Sangat Mudah
   11            11           28           93,33  Sangat Mudah
   12            12           27           90,00  Sangat Mudah
   13            13           30          100,00  Sangat Mudah
   14            14           29           96,67  Sangat Mudah
   15            15           23           76,67  Mudah
  
```

Tabel 8. Hasil Tingkat Kesukaran Butir Soal

Kategori Soal	Jumlah	Persentase%	Nomor Soal
Mudah	3	20%	2, 4, dan 15
Sangat Mudah	12	80%	1, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, dan 14

Tabel 8 menunjukkan hasil uji coba tes pilihan ganda, ada 3 soal yang merupakan tingkatan mudah dengan presentase 20% dan 12 soal dengan tingkatan kesukaran nya sangat mudah dengan presentase 80%. Maka dari itu, nomor 1 hingga 15 memiliki tingkatan kesukaran mudah semua. Dengan indeks 0,71 – 1,00.

Berdasarkan data yang tertera di atas, dapat disimpulkan bahwa soal ini terlalu mudah dan membutuhkan perbaikan menjadi soal yang lebih baik

dikarenakan soal harus mempunyai tingkat kesulitan yang seimbang. Artinya, tidak hanya beberapa dari soal-soal yang mudah atau sulit saja. Adanya variasi tingkat kesulitan, seperti 35% soal mudah, 50% soal sedang, dan 15% soal sulit, akan membuat ujian lebih efektif dalam mengukur kemampuan siswa. Jika di suatu soal yang terlalu sederhana atau mudah tidak akan merangsang siswa agar berpikir lebih dalam, sedangkan soal yang terlalu sulit dapat menurunkan motivasi belajar siswa. Oleh karena itu, tingkat kesulitan soal harus disesuaikan dengan tahap perkembangan agar tujuan evaluasi pembelajaran dapat dicapai oleh siswa. Untuk membuat soal sulit dipahami oleh siswa, evaluasi dapat dilakukan. Berdasarkan masalah tersebut, siswa dapat memilih kembali pertanyaan mana yang sesuai atau meneliti bagian dari soal yang paling menantang untuk disesuaikan dengan tingkat kesulitan yang diinginkan (Irawati. R. dkk., 2020).

4. Daya Pembeda Butir Soal

Software Anates digunakan agar bisa analisis butir soal yang menghasilkan nilai daya beda yang dikategorikan menjadi empat level yaitu, sangat baik, baik, cukup, dan jelek. Nilai daya beda ini bisa bernilai positif atau negatif, yang mengindikasikan seberapa efektif elemen soal dengan memisahkan siswa yang dapat menerapkan materi dengan siswa yang belum dapat menerapkan materi. Berikut adalah bukti hasil daya pembedanya :

Gambar 3. Daya Pembeda Butir Soal

```

DAYA PEMBEDA
=====

Jumlah Subyek= 30
Klp atas/bawah(n)= 8
Butir Soal= 15
Nama berkas: C:\USERS\ASUS\ONEDRIVE\KULIAH\MATA KULIAH SEMESTER 3\EVALUASI BE
    
```

No Butir Baru	No Butir Asli	Kel. Atas	Kel. Bawah	Beda	Indeks DP (%)
1	1	8	6	2	25,00
2	2	8	3	5	62,50
3	3	8	5	3	37,50
4	4	8	2	6	75,00
5	5	8	6	2	25,00
6	6	8	7	1	12,50
7	7	8	8	0	0,00
8	8	8	7	1	12,50
9	9	8	6	2	25,00
10	10	8	7	1	12,50
11	11	8	8	0	0,00
12	12	8	7	1	12,50
13	13	8	8	0	0,00
14	14	8	7	1	12,50
15	15	7	3	4	50,00

Tabel 9. Hasil Daya Pembeda

Kategori Soal	Jumlah	Persentase%	Nomor Soal
Buruk	8	53,3%	6, 7, 8, 10, 11, 12, 13, dan 14
Cukup	4	26,6%	1, 3, 5, dan 9
Baik	2	13,3%	2 dan 15
Sangat Bak	1	6,6%	4

Hasil pada analisis kemampuan daya pembeda butir soal yang menggunakan *software* Anates menunjukkan bahwa 8 butir soal ini termasuk kategori buruk pada nomor soal 6, 7, 8, 10, 11, 12, 13, dan 14 yang berada di batasan 0,00 – 0,20 dengan persentase 53,3%. Terdapat 4 butir soal termasuk kategori cukup pada nomer soal 1, 3, 5, dan 9 yang berada di batasan 0,20 – 0,40 dengan presentase 26,6%. Terdapat 2 butir soal ini termasuk kategori baik pada nomer soal 2 dan 15 yang berada di batasan 0,40 – 0,70. Dan yang terakhir terdapat 1 butir soal termasuk kategori sangat baik pada nomer soal 4 berada di batasan 0,70 – 1,00 dengan presentase 6,6%.

5. Analisis Kelompok Unggul dan Asor

Gambar 4. Hasil Pengolaan Kelompok Unggul dan Asor

Kelompok Unggul																		
No. Urut	No Subyek	Kode>Nama Subyek	Skor	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	2	CYNTIA AYU VIVIA	15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	4	AQILA SALSABILA	15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	5	JUNI TRIWAHYUNI	15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	7	ANDINI CAHYA VITASARI	15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	11	IRDA SENIORITA	15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	14	KHOIRUN NISA'	15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	29	RAFIF HAYDAR GUNAWAN	15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	1	KHOLISATUN NISA	14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	-

Kelompok Asor																		
No. Urut	No Subyek	Kode>Nama Subyek	Skor	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	22	LATIFA NURUL HIDAYATI	13	1	-	1	-	1	1	1	1	1	1	1	1	1	1	1
2	27	NENSYA RACHMA SALSABILA	13	1	1	1	1	1	1	1	1	-	1	1	1	1	1	-
3	18	SASKIA N.N	12	1	-	1	-	1	1	1	1	1	1	1	1	1	1	-
4	19	RAHMA ALWI	12	-	1	-	-	1	1	1	1	1	1	1	1	1	1	1
5	20	THELAN ANDJANI	12	1	1	1	1	-	1	1	-	1	1	1	1	1	1	-
6	24	ERLINDA NUR FADILLAH	12	1	-	1	-	1	1	1	1	1	1	1	1	1	1	-
7	30	ZEBINA FATATURROHMAH	10	1	-	-	-	-	1	1	1	1	1	1	1	1	1	-
8	28	DZULTRIFAN SYA'BANI RAJHAN	6	-	-	-	-	1	-	1	1	-	-	1	-	1	-	1

Berdasarkan hasil pengolaan terhadap kelompok unggul dan asor butir soal dengan *software* Anates menunjkan bahwa hasil dari 27% dari 30 peserta tes uji

coba soal Ada delapan peserta ujian dalam kelompok unggul (kelompok atas) menjawab soal dengan jumlah benar tertinggi, yaitu Cyntia, Aqila, Juni, Andini, Irda, Khoirun, Rafif, dan Kholisatun. Sedangkan hasil dari 27% dari 30 peserta tes uji coba soal. Dalam kelompok bawah, terdapat delapan peserta ujian yang memberikan jawaban untuk soal dengan angka terendah, yaitu Latifa, Nensya, Saskia, Rahma, Thelan, Erlinda, Zebina, dan Zul’irfan.

Software Anates tidak hanya mengklasifikasikan siswa menjadi dua kelompok kemampuan ekstrem (tinggi dan rendah), tetapi juga mampu mengidentifikasi siswa dengan kemampuan yang berada di sekitar nilai KKM. Siswa dengan kemampuan sedang ini memiliki prestasi yang relatif dekat dengan standar kompetensi minimum. Soal-soal yang dihasilkan oleh *software* Anates dan dikategorikan sebagai sangat baik, baik, atau cukup dapat disimpan dan digunakan kembali dalam penilaian berikutnya. Soal-soal yang dinilai kurang baik perlu diperbaiki atau diganti agar meningkatkan kualitas penilaian.

6. Korelasi Skor Butir dengan Skor Total

Yang di maksud korelasi antara butir soal dengan skor total merupakan indikator penting dalam analisis tes. Korelasi ini menggambarkan sejauh mana jawaban pada butir soal terkait dengan pencapaian peserta secara keseluruhan (Aurelia, T. R. dkk., 2024).

Gambar 5. Hasil Korelasi Skor Butir dengan Skor Total

```

KORELASI SKOR BUTIR DG SKOR TOTAL
=====
Jumlah Subyek= 30
Butir Soal= 15
Nama berkas: C:\USERS\ASUS\ONEDRIVE\KULIAH\MATA KULIAH SEMESTER 3
    
```

No Butir Baru	No Butir Asli	Korelasi	Signifikansi
1	1	0,649	Sangat Signifikan
2	2	0,658	Sangat Signifikan
3	3	0,748	Sangat Signifikan
4	4	0,584	Signifikan
5	5	0,244	-
6	6	0,377	-
7	7	NAN	NAN
8	8	0,080	-
9	9	0,501	Signifikan
10	10	0,439	-
11	11	-0,094	-
12	12	0,439	-
13	13	NAN	NAN
14	14	0,761	Sangat Signifikan
15	15	0,332	-

Tabel 10. Hasil Korelasi Skor Butir dengan Skor Total

Kategori Soal	Jumlah	Persentase%	Nomor Soal
Validasi Tinggi	4	26,6%	1, 2, 3, dan 14
Validasi Cukup	2	13,3%	4 dan 9
Validasi Rendah	5	33,3%	5, 6, 10, 12, dan 15
Validasi sangat rendah	4	26,6%	7, 8, 11, dan 13

Tabel 10 menunjukkan hasil pengolahan terhadap korelasi antara skor masing-masing butir soal dan skor totalnya dengan *software* Anates, mengingat bahwa terdapat 4 butir soal pada 1, 2, 3, dan 14 yang termasuk validasi tinggi dengan persentase 26,6% . Terdapat juga 2 butir soal pada nomor soal 4 dan 9 yang termasuk validasi cukup dengan persentase 13,3%. Terdapat 5 butir soal pada nomor soal 5, 6, 10, 12, dan 15 yang termasuk validasi rendah dengan persentase 33,3%. Yang terakhir, terdapat 4 butir soal pada nomor 7,8,11,dan 13 yang termasuk validasi sangat rendah dengan persentase 26% .

Berdasarkan hasil data diatas dapat disimpulkan bahwa terdapat 4 soal valid dan 11 soal membutuhkan perbaikan karena jika soal yang digunakan itu tidak valid, maka hasil penilaian yang diperoleh itu tidak akan mencerminkan kemampuan nyatanya atau sebenarnya dari peserta didik. Soal tidak valid, berarti soal tersebut tidak melakukan pengukuran yang seharusnya dilakukan. Validitas dan reliabilitas merupakan dua konsep yang saling berkaitan dalam penilaian. Soal yang reliabel cenderung menghasilkan hasil yang valid. Dengan kata lain, suatu soal kategori baik harus memenuhi kedua hal ini. Soal yang sudah terbukti valid dan reliabel dapat disimpan untuk digunakan kembali dalam penilaian berikutnya. Sebaliknya, soal yang tidak valid perlu dilakukan perbaikan atau diganti agar sesuai dengan standar kriteria pembuatan soal yang baik. Guru dapat melibatkan para ahli agar bisa membantu memperbaiki soal yang tidak valid.

7. Kualitas Pengecoh

Hasil pemeriksaan HOTS menggunakan *software* Anates dipecah menjadi 5 (lima) tingkatan berbeda, yaitu sangat buruk, buruk, kurang baik, baik, dan sangat baik. Pola sebaran jawaban dapat digunakan untuk mengukur efektivitas pengecoh.

Pengecoh bertujuan untuk mengelabui siswa yang tidak seberapa tahu dalam mata pelajaran yang mereka pelajari. (Rahayu, L. P. dkk., 2024)

Gambar 6. Hasil Kualitas Pengecoh

```

KUALITAS PENGECOH
=====
Jumlah Subyek= 30
Butir Soal= 15
Nama berkas: C:\USERS\ASUS\ONEDRIVE\KULIAH\MATA KULIAH SEMESTER 3\EVALUA.

No Butir Baru  No Butir Asli  a  b  c  d  e  *
1 1 1-- 1-- 28** 0-- 0-- 0
2 2 0-- 6-- 24** 0-- 0-- 0
3 3 2-- 27** 0-- 1+ 0-- 0
4 4 2++ 22** 2++ 4-- 0-- 0
5 5 2-- 26** 1++ 1++ 0-- 0
6 6 1+ 1+ 27** 1+ 0-- 0
7 7 0 30** 0 0 0 0
8 8 26** 2-- 1++ 1++ 0-- 0
9 9 1+ 1+ 27** 0-- 1+ 0
10 10 27** 2-- 0-- 0-- 1+ 0
11 11 1-- 0-- 28** 1-- 0-- 0
12 12 27** 2-- 0-- 1+ 0-- 0
13 13 0 30** 0 0 0 0
14 14 0-- 0-- 1-- 29** 0-- 0
15 15 23** 3- 2++ 0-- 2++ 0
    
```

Tabel 11. Hasil Kualitas Pengecoh

Kategori Soal	Jumlah	Persentase%	Nomor Soal
Sangat Buruk	6	40%	1, 2, 7, 11, 13, dan 14
Buruk	3	20%	3, 10, dan 12
Kurang Baik	4	26,6%	4, 5, 8, dan 15
Baik	2	13,3%	6 dan 9
Sangat Baik	0	0%	0

Tabel 11 menunjukkan hasil pengolahan terhadap kualitas pengecoh, pengecoh dapat dikatakan bisa melakukan tugasnya dengan baik jika mereka menarik, sehingga kekurangan pemahaman konsep sering kali membuat peserta didik ragu-ragu dan cenderung memilih opsi yang keliru sebagai jawaban yang benar.

Setiap soal memiliki skala dari sangat baik hingga sangat buruk untuk mengukur efektivitas pengecoh.

- 1) Sebuah soal dikatakan sangat efektif jika keempat opsi jawaban yang salah berfungsi sebagai pengalih perhatian yang baik.
- 2) Sebuah soal dikatakan efektif jika tiga dari empat opsi jawaban yang salah berfungsi sebagai pengalih perhatian yang baik.
- 3) Sebuah soal dikatakan kurang efektif jika hanya dua opsi jawaban yang salah berfungsi sebagai pengalih perhatian yang baik.

- 4) Sebuah soal dikatakan tidak efektif jika hanya satu opsi jawaban yang salah berfungsi sebagai pengalih perhatian yang baik.
- 5) Sebuah soal dikatakan sangat tidak efektif jika tidak ada satupun opsi jawaban yang salah berfungsi sebagai pengalih perhatian.

Berdasarkan analisis kualitas butir soal yang mengecoh, yaitu terdapat 6 butir soal pada nomor 1, 2, 7, 11, 13, dan 14 yang memiliki Kriteria Efektivitas Pengecoh butir soal Sangat Buruk dengan presentase 40% karena keempat pengecohnya tidak berfungsi. Lalu terdapat 3 butir soal pada nomor 3, 10, dan 12 yang memiliki Kriteria Efektivitas Pengecoh butir soal Buruk dengan presentase 20% karena hanya terdapat satu pengecoh saja yang berfungsi. Terdapat 4 butir soal pada nomor 4, 5, 8, dan 15 yang memiliki Kriteria Efektivitas Pengecoh butir soal Kurang Baik dengan presentase 26,6% karena hanya dua pengecoh yang berfungsi. Kemudian, terdapat 2 butir soal pada nomor 6 dan 9 yang memiliki Kriteria Efektivitas Pengecoh butir soal Baik dengan presentase 13,3% dengan memiliki tiga pengecoh yang berfungsi. Dari kelima belas butir soal, tidak terdapat Kriteria Efektivitas Pengecoh Detail Pertanyaan yang Dikatakan Sangat Baik karena tidak terdapat keempat pengecoh yang berfungsi, oleh karena itu, setiap tugas yang presentase indeks pengecohnya dikategorikan sebagai kurang baik, buruk, atau sangat buruk, memerlukan perbaikan.

8. Rekap Analisis Butir Soal

Gambar 7. Hasil Rekap Analisis Butir Soal

Preview

Preview Kembali Ke Menu Sebelumnya Cetak ke file Cetak ke printer

REKAP ANALISIS BUTIR
 =====

Rata2= 13,37
 Simpang Baku= 1,83
 KorelasiXI= 0,48
 Reliabilitas Tes= 0,65
 Butir Soal= 15
 Jumlah Subyek= 30
 Nama berkas: C:\USERS\ASUS\ONEDRIVE\KULIAH\MATA KULIAH SEMESTER 3\EVALUASI BELAJAR DAN PEMBELAJARAN\TUGAS\HASIL UJI COBA ANATES MDP.ANA

Btr Baru	Btr Asli	D.Pembeda(%)	T. Kesukaran	Korelasi	Sign. Korelasi
1	1	25,00	Sangat Mudah	0,649	Sangat Signifikan
2	2	62,50	Mudah	0,658	Sangat Signifikan
3	3	37,50	Sangat Mudah	0,748	Sangat Signifikan
4	4	75,00	Mudah	0,584	Signifikan
5	5	25,00	Sangat Mudah	0,244	-
6	6	12,50	Sangat Mudah	0,377	-
7	7	0,00	Sangat Mudah	NAN	NAN
8	8	12,50	Sangat Mudah	0,080	-
9	9	25,00	Sangat Mudah	0,501	Signifikan
10	10	12,50	Sangat Mudah	0,439	-
11	11	0,00	Sangat Mudah	-0,094	-
12	12	12,50	Sangat Mudah	0,439	-
13	13	0,00	Sangat Mudah	NAN	NAN
14	14	12,50	Sangat Mudah	0,761	Sangat Signifikan
15	15	50,00	Mudah	0,332	-

Tabel 12. Hasil Rekap Analisis Butir Soal

Kategori Soal	Jumlah	Persentase%	Nomor Soal
Soal Diperbaiki	8	53,3%	6, 7, 8, 10, 11, 12, 13, dan 14
Soal Digunakan dan Diperbaiki	4	26,6%	1, 3, 5, dan 9
Soal Digunakan	3	20%	2, 4, dan 15

Berdasarkan hasil rekap analisis butir soal yang ditunjukkan dalam tabel 12 diketahui bahwa terdapat 8 butir soal pada soal nomor 6, 7, 8, 10, 11, 12, 13, dan 14 yang harus diperbaiki karena kurangnya daya pembeda, tingkat kesukaran, dan kualitas pengecoh, dengan persentase 53,3%. Kemudian terdapat 4 butir soal pada soal nomor 1, 3, 5, dan 9 yang bisa digunakan namun masih harus dilakukan perbaikan juga, dikarenakan masih terdapat kekurangan yang mungkin ada pada daya pembeda, tingkat kesukaran, ataupun pada kualitas pengecoh. Dan yang terakhir, terdapat 3 butir soal pada nomor 2, 4, dan 15 yang bisa digunakan, dimana ketiga bagian soal tersebut dapat langsung digunakan sebagai soal yang dapat diujikan kepada peserta didik.

KESIMPULAN DAN SARAN

Anates memiliki kemampuan untuk melakukan evaluasi butir soal pilihan ganda dan deskriptif sehingga memberikan reliabilitas, daya pembeda, kesulitan, dan

efektivitas pengecoh soal, yang menjadikannya solusi (Naibaho. M. A. U. dkk., 2021). Elviana (2020) mengutarakan jika Daya Pembeda adalah indeks yang dapat menunjukkan seberapa baik siswa melakukan tugas-tugas tertentu untuk bisa membedakan kelompok atas (siswa yang unggul, yang memiliki nilai tes yang tinggi) dari kelompok bawah (siswa yang asor, yang memiliki nilai tes yang rendah). Pada proses validasi dan juga reliabilitas banyak dipengaruhi faktor terkait hasil validitas dan hasil reliabilitas. Hasil validitas dan reliabilitas akan berpengaruh pada kualitas dan kuantitas yang dihasilkan pada sebuah penelitian (Fadli, R., dkk. 2023). Menurut hasil analisis, 15 soal pilihan ganda yang memuat Elemen Praktikum Perusahaan Jasa, Dagang, dan Manufaktur yang menggunakan *Software Anates 4.0.9*, kualitas soal pilihan ganda yang memuat Elemen Praktikum Perusahaan Jasa, Dagang, dan Manufaktur belum baik. Hasil dari analisis butir soal yaitu lima belas soal yang diberikan kepada 30 siswa kelas XI menunjukkan bahwa, terdapat 4 butir soal yang dapat digunakan namun masih harus diperbaiki dan terdapat 8 butir soal harus melakukan perbaikan. Sehingga hanya tersisa 3 butir soal dari 15 soal yang dapat digunakan untuk bahan uji coba peserta didik. Diharapkan pembaca bisa memahami kualitas soal-soal yang terbuat dari pembahasan materi di atas, sehingga mereka dapat memperbaiki soal-soal berikutnya. Kemudian, penyusun soal juga perlu untuk mengikuti pelatihan – pelatihan pembuatan soal supaya dapat membuat soal yang lebih baik lagi kedepannya.

DAFTAR REFERENSI

- Ambarawati, Y.F. & Ismiyati, I. (2021). Analisis butir soal pilihan ganda ulangan akhir semester genap mata pelajaran kearsipan. *Measurement in Educational Research*, 1(2). <http://dx.doi.org/10.33292/meter.v1i2.144>
- Amelia, N., Paridjo, & Sina, I. (2021). Analisis butir soal penilaian akhir tahun mata pelajaran matematika menggunakan software anates. *AKSIOMA: Jurnal Matematika Dan Pendidikan Matematika*, 12(3), 447–456. <https://doi.org/10.26877/aks.v12i3.8356>
- Anwar, B., Buku Master Book Kimia SMA-MA/SMK-MAK Kelas X, XI, XII, 2021, Bandung, Yrama Widya, ISBN : 978-602-374-952-2
- Ashari, S. D. (2021). Tingkat Kesukaran, Daya Pembeda Dan Efektivitas Pengecoh Soal Ulangan Akhir Semester Pjok Kelas Iv Dan V Sd Negeri Ngombol, Purworejo Tahun Ajaran 2020/2021. *Angewandte Chemie International Edition*, 6(11), 951–952., 2013–2015. <http://eprints.uny.ac.id/id/eprint/71423>
- Aurelia, T. R., Agnola, E. B. Y., Hakim, L., & Pratiwi, V. (2024). ANALISIS EVALUASI BUTIR SOAL MATA PELAJARAN KOMPUTER AKUNTANSI DENGAN SOFTWARE ANATES OLEH PESERTA DIDIK SMK KETINTANG. *Jurnal Ilmiah Kajian Multidisipliner*, 8(11). <https://oaj.jurnalhst.com/index.php/jikm/article/view/6317>
- Dewi P. Elvia R. Elvinawati. 2021. PENGEMBANGAN BUTIR SOAL HOTS UNTUK MENGUJI KEMAMPUAN BERPIKIR TINGKAT TINGGI SISWA DI MA NEGERI 2 KOTA BENGKULU. *Jurnal Pendidikan dan Ilmu Kimia*, 5(2): 141 – 148. ISSN 2615-2819. <https://doi.org/10.33369/atp.v5i2.17119>
- Elviana. (2020). Analisis Butir Soal Evaluasi Pembelajaran Pendidikan Agama Islam Menggunakan Program Anates. *Jurnal Mudarrisuna*, 10(2), 58-74. <http://dx.doi.org/10.22373/jm.v10i2.7839>
- Fadli, R., Hidayati, S., Cholifah, M., Siroj, R. A., & Afgani, M. W. (2023). Validitas dan Reliabilitas pada Penelitian Motivasi Belajar Pendidikan Agama Islam Menggunakan Product Moment. *JiIP-Jurnal Ilmiah Ilmu Pendidikan*, 6(3), 1734-1739. <http://dx.doi.org/10.54371/jiip.v6i3.1419>
- Fiska, Joelan M., Yunin H., Nur Qomaria., dan Wiwin Puspita H. (2021). Analisis Butir Soal Ulangan Harian IPA Menggunakan Software Anates pada Pendekatan

Teori Tes. *Jurnal Natural Science Educational Research*.4(1)
<https://doi.org/10.21107/nser.v4i1.8133>

- Fietri, W. A., Lufri, Syamzurizal, & Zulyusri. (2021). Analisis Butir Soal Biologi Kelas VIII Madrasah Tsanawiyah Negeri 6 Kerinci. *Jurnal Pendidikan Biologi Undiksha*, 8(2), 50–60. <https://doi.org/10.23887/jjpb.v8i2>
- Gradini,E., Menilik Konsep Kemampuan BerpikirTingkat Tinggi (Higher Order Thinking Skills) Dalam Pembelajaran Matematika,Jurnal Numeracy, 2019, 6(2): 189-203. <https://doi.org/10.46244/numeracy.v6i2.475>
- Irawati, R., Ekawati, E. Y., & Budiawanti, S. (2020). Analisis Butir Soal Ujian Akhir Semester Gasal Menggunakan Program Anbuso di SMA Negeri 1 Boyolali Tahun Ajaran 2019/2020. *Jurnal Materi Dan Pembelajaran Fisika*, 10(1), 11. <https://doi.org/10.20961/jmpf.v10i1.42084>
- Kaka L. Bano O V. Dkk. 2024. EFEKTIVITAS ANALISIS BUTIR SOAL PILIHAN GANDA MENGGUNAKAN APLIKASI ANATES DI SMPN 2 KANATANG. *Jurnal Inovasi Penelitian*. ISSN 2722-9475. Vol. 4 No. 2. <https://doi.org/10.47492/jip.v4i9.3124>
- Lestari D L. Yumma Z R. Dkk. 2024. Analisis Kualitas Butir Soal HOTS pada Mata Pelajaran Biologi yang Diberikan pada Siswa Kelas 12 dan Dievaluasi Menggunakan Aplikasi Anates. *Journal of Basic Educational Studies*. ISSN 2775-7269. Vol 4 No 2. <https://doi.org/10.47467/edu.v4i2.1935>
- Muchlizani, N. A.,Mania, S., & Rasyid, M. N. A. 2023. Analisis Kualitas Butir Soal Ujian Akhir Semester Mata Pelajaran Akidah Akhlak Kelas V Mi Radhiatul Adawiyah Makassar. *Jurnal Inspiratif Pendidikan*, 12(1), 224–240. <https://doi.org/10.24252/ip.v12i1.32040>
- Naibaho, M. A. U., Taroreh, M. L. G., & Kapantow, G. H. M. (2021). Kajian Keberlanjutan Program Yayasan Pembangunan Berkelanjutan Sulawesi Utara (Ypbsu) Pasca Penutupan Tambang Pt. Newmont Minahasa Raya (Pt. Nmr) Di Kecamatan Ratatotok Minahasa Tenggara. *Agri-Sosioekonomi*, 17(2), 177. <https://doi.org/10.35791/agrsosek.17.2.2021.33834>
- Pradita, E., Megawanti, P., & Yulianingsih, Y. (2023). Analisis Tingkat Kesukaran, Daya Pembeda, dan Fungsi Distraktor PTS Matematika SMPN Jakarta. *Himpunan: Jurnal Ilmiah Mahasiswa Pendidikan Matematika*, 3(1), 109-118. <https://jim.unindra.ac.id/index.php/himpunan/article/view/8645>

- Pramarth, I. N. B., Rathintara, I. A. L., & Astapa, I. G. (2023). Pengembangan Instrumen Tes Kemampuan Berpikir Kritis Berbasis Higher Order Thinking Skills (HOTS). *Journal on Education*, 6(1), 680-686.
- Rahayu, L. P., & Sukenti, D. (2024). Kualitas Soal Bahasa Indonesia Kelas XI SMAN 2 Bangko Pusako: Analisis Butir Soal. *Jurnal Onoma: Pendidikan, Bahasa, dan Sastra*, 10(4), 3755-3762. <https://www.e-journal.my.id/onoma/article/view/4481>
- Ramadhan, W., Malahati, F., Romadhon, K., & Ramadhan, S. (2023). Analisis Butir Soal Tipe Multiple Choice Questions pada Penilaian Harian Sekolah Dasar. *Tarbiyah Wa Ta'lim: Jurnal Penelitian Pendidikan Dan Pembelajaran*, 10(2), 93–105. <https://doi.org/10.21093/twt.v10i2.6155>
- Ratih, P. D., Pristanti, V. A., Hakim, L., & Pratiwi, V. (2024). Analisis Butir Soal Pilihan Ganda Berbasis Higher Order Thinking Skills Menggunakan Software Anates Versi 4.0 Pada Fase E Elemen Etika Profesi Di Bidang Akuntansi Dan Keuangan Lembaga Siswa Smk Jurusan Akuntansi Kelas X. *Jurnal Ilmiah Kajian Multidisipliner*, 8(11). <https://oaj.jurnalhst.com/index.php/jikm/issue/view/366>
- Saputri, H. A. S., & Larasati, N. J. (2023). ANALISIS INSTRUMEN ASSESMEN: VALIDITAS, RELIABILITAS, TINGKAT KESUKARAN DAN DAYA BEDA BUTIR SOAL. *Didaktik: Jurnal Ilmiah PGSD STKIP Subang*, 9(5), 2986-2995. <https://doi.org/10.36989/didaktik.v9i5.2268>
- Saraswati, P.M.S., dan Gusti Ngurah Sastra Agustika, Kemampuan Berpikir Tingkat Tinggi Dalam Menyelesaikan Soal HOTS Mata Pelajaran Matematika, *Jurnal Ilmiah Sekolah Dasar*, 2020, 4(2) : 257-269. <https://doi.org/10.23887/jisd.v4i2.25336>
- Widhiyani, I. A. N. T., I. N. Sukajaya, dan G. Suweken., Pengembangan Soal Higher Order Thinking Skills Untuk Pengkategorian Kemampuan Pemecahan Masalah Geometri Siswa SMP, *Jurnal Pendidikan dan Pembelajaran Matematika Indonesia*, 2019, 8(2): 161 – 170. <https://doi.org/10.23887/jppm.v8i2.2854>
- Yusuf W F. 2024. ANALISIS BUTIR SOAL ASESMEN SUMATIF BIOLOGI MATERI PERUBAHAN LINGKUNGAN DENGAN MENGGUNAKAN ANATES PADA KELAS X SMA. *Jurnal Inovasi Penelitian Pendidikan dan Pembelajaran*. ISSN : 2777-0575. Vol. 4 No. 1. <https://doi.org/10.51878/learning.v4i1.2777>

