

Nama : Alfiya Nadhira Syifa
NPM : 2413031037
Kelas : 2024 B
Mata Kuliah : Evaluasi Pembelajaran Ekonomi
Dosen Pengampu : Prof. Dr. Herpratiwi, M.Pd
Dr. Pujiati, S.pd., M.Pd.
Fanni Rahmawati, S.pd., M.Pd.

ANALISIS KUALITAS TES DAN BUTIR SOAL

Informasi Pelaksanaan

Nama Sekolah : SMA Global Madani
Mata Pelajaran : Ekonomi
Materi : APBN dan APBD
Kelas : XI.4
Waktu Pelaksanaan : 11 Mei 2026
Jumlah Responden : 18 siswa
Alat Uji Instrumen : <https://forms.gle/xaEsHB7ex8NkHd9o> (link gform)

ANALISIS ASPEK PENGETAHUAN

1. Uji Validitas

Uji validitas bertujuan untuk mengetahui sejauh mana instrumen tes mampu mengukur apa yang seharusnya diukur. Pengujian validitas butir soal umumnya dilakukan dengan mengkorelasikan skor setiap butir dengan skor total menggunakan uji korelasi Product Moment Pearson.

Namun, pada penelitian ini uji validitas butir soal tidak dapat dilakukan secara optimal karena jumlah responden yang mengerjakan instrumen hanya 18 siswa. Jumlah tersebut belum memenuhi jumlah sampel yang direkomendasikan untuk analisis validitas butir. Dalam analisis statistik, pengujian validitas umumnya memerlukan jumlah responden yang lebih besar agar hasil korelasi yang diperoleh stabil dan dapat mewakili populasi. Beberapa referensi metodologi penelitian menyarankan jumlah responden minimal 30 orang, bahkan lebih baik apabila mencapai 5-10 kali jumlah butir soal yang diuji.

Dengan jumlah responden yang terbatas, nilai korelasi yang dihasilkan berpotensi kurang stabil dan kurang mencerminkan kualitas butir soal yang sebenarnya. Oleh

karena itu, hasil uji validitas tidak dapat dijadikan dasar yang kuat untuk menyimpulkan valid atau tidaknya setiap butir soal dalam instrumen ini.

2. Uji Reliabilitas

Uji reliabilitas bertujuan untuk mengetahui tingkat konsistensi instrumen dalam mengukur suatu variabel. Instrumen yang reliabel akan memberikan hasil yang relatif konsisten apabila digunakan pada kelompok responden yang memiliki karakteristik serupa. Pengujian reliabilitas umumnya dilakukan menggunakan koefisien reliabilitas, seperti Cronbach's Alpha untuk instrumen dengan skor bertingkat atau KR-20/KR-21 untuk soal objektif.

Pada observasi ini, uji reliabilitas tidak dilakukan karena jumlah responden yang mengerjakan instrumen hanya 18 siswa. Jumlah responden tersebut dinilai belum memadai untuk menghasilkan estimasi koefisien reliabilitas yang stabil dan akurat. Dalam analisis instrumen, pengujian reliabilitas memerlukan jumlah sampel yang cukup agar hasil perhitungan dapat menggambarkan konsistensi instrumen secara lebih representatif.

Apabila uji reliabilitas dilakukan dengan jumlah responden yang terlalu sedikit, nilai koefisien reliabilitas yang diperoleh berpotensi mengalami fluktuasi yang tinggi sehingga kurang dapat dijadikan dasar dalam menilai tingkat keandalan instrumen. Oleh karena itu, hasil pengujian reliabilitas dengan jumlah sampel yang terbatas dikhawatirkan tidak mencerminkan kondisi instrumen yang sebenarnya.

3. Tingkat Kesukaran (P)

Tingkat Kesukaran adalah ukuran yang menunjukkan mudah atau sulitnya suatu butir soal bagi peserta didik. Analisis tingkat kesukaran dilakukan untuk mengetahui proporsi peserta didik yang dapat menjawab suatu soal dengan benar.

Keterangan:

Pilihan Ganda

- 1 untuk jawaban benar
- 0 untuk jawaban salah

Soal Uraian

- Skor dihitung berdasarkan rubrik penilaian

A. Pilihan Ganda

Rumus Indeks Kesukaran:

$$P = \frac{\sum B}{N}$$

Keterangan:

- P = angka indeks kesukaran item
- $\sum B$ = jumlah siswa yang menjawab benar

- N = jumlah siswa yang mengikuti tes hasil

Indeks Tingkat Kesukaran:

Indeks Tingkat Kesukaran	Klasifikasi
0,00-0,30	Sukar
0,31-0,70	Sedang
0,71-1.00	Mudah

Tabel Hasil Uji Skor Pengetahuan Siswa Soal Pilihan Ganda

No	Nama	Nomor Soal																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	Ahmad Faiz Raziq	1	1	1	1	1	1	1	1	0	1	1	0	0	1	1	1	1	1	0	1
2	Amayra Arkanaya	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1
3	Aura Sabrina Putri	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
4	Bagas Putra Kholik P.	1	1	1	0	1	1	1	1	1	1	0	1	1	0	0	1	1	1	1	1
5	Fayza Kulla Che A.	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	Luthfi Arya Satia	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
7	M. Datik Pratama D. P	1	1	1	1	1	1	0	0	1	1	0	1	1	1	0	1	1	0	0	1
8	M. Zaky Azwa Al Z.	1	1	1	1	1	0	1	1	1	1	0	0	1	1	1	1	1	1	0	1
9	Muhammad Akbar M.	0	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1
10	Muhammad Fahri A.	0	1	1	1	0	1	1	0	0	1	0	1	1	1	0	1	1	1	0	0
11	Muhammad Rakha K.	1	0	1	1	0	1	1	1	1	0	0	1	0	1	1	0	0	1	0	1
12	Nuryan Husnul K	1	1	1	1	0	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1
13	Raina Sava R.	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
14	Ridwan Prima Satifa	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	Riezka Nadya Putri	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	Rio Dwi Arviansyah	1	1	1	1	0	1	0	1	1	1	0	1	1	1	0	0	1	1	1	1
17	Sulthan Laksamana	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	1
18	Tafakanara Alsajada	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0

Tabel Hasil Interpretasi Derajat Kesukaran Pilihan Ganda

No	B	N	P	Kategori
1	16	20	0,80	Mudah
2	16	20	0,80	Mudah
3	18	20	0,90	Mudah
4	16	20	0,80	Mudah
5	12	20	0,60	Sedang
6	17	20	0,85	Mudah
7	15	20	0,75	Mudah
8	12	20	0,60	Sedang
9	14	20	0,70	Sedang

10	15	20	0,75	Mudah
11	11	20	0,55	Sedang
12	15	20	0,75	Mudah
13	16	20	0,80	Mudah
14	17	20	0,85	Mudah
15	14	20	0,70	Sedang
16	16	20	0,80	Mudah
17	17	20	0,85	Mudah
18	16	20	0,80	Mudah
19	12	20	0,60	Sedang
20	16	20	0,80	Mudah

Hasil analisis menunjukkan bahwa instrumen tes ini didominasi oleh kategori Mudah sebanyak 14 butir soal (70%) dan kategori Sedang sebanyak 6 butir soal (30%), tanpa adanya soal dalam kategori Sukar. Secara keseluruhan, kualitas tes ini cenderung memiliki tingkat kesulitan yang rendah bagi siswa karena mayoritas soal memiliki indeks di atas 0,70.

B. Esai

Rumus Indeks Kesukaran Soal Uraian:

$$P = \frac{\bar{x}}{SMI}$$

- P = Indeks Kesukaran
- \bar{x} = Rata-rata skor yang diperoleh siswa
- SMI = Skor Maksimum Ideal

No Soal	SMI	Rata-rata Skor (\bar{x})	Indeks Kesukaran (P)	Keterangan
1	10	8,5	0,85	Mudah
2	10	7,2	0,72	Mudah
3	15	9,0	0,60	Sedang
4	15	6,5	0,43	Sedang
5	20	17,0	0,85	Mudah

Kesimpulan Akhir

Berdasarkan analisis pada bagian Pilihan Ganda dan Esai, instrumen tes secara umum memiliki tingkat kesukaran yang didominasi oleh kategori Mudah dan Sedang. Pada Pilihan Ganda, 70% soal berkategori Mudah dan 30% Sedang. Sementara pada bagian Esai, terdapat variasi yang lebih baik dengan 3 soal kategori Mudah dan 2 soal kategori Sedang. Tidak ditemukan soal kategori Sukar dalam instrumen ini, sehingga disarankan untuk menambah butir soal dengan tingkat tantangan lebih tinggi pada evaluasi berikutnya.

4. Daya Pembeda

A. Pilihan Ganda

Rumus: $D = PA - PB$

Keterangan:

- D = Daya Pembeda
- PA = Proporsi kelompok atas yang menjawab benar
- PB = Proporsi kelompok bawah yang menjawab benar
- JA = Banyak peserta kelompok atas
- JB = Banyak peserta kelompok bawah
- BA = Banyak peserta kelompok atas yang menjawab benar
- BB = Banyak peserta kelompok bawah yang menjawab benar

Kelompok Kelas Atas (JA)

Diambil dari barisan siswa dengan skor tertinggi.

No	Nama	Skor	Soal Salah
1	Ridwan Prima Satifa	100	-
2	Amayra Arkanaya	95	1
3	Fayza Kulla Che Azima	95	1
4	Luthfi Arya Satia	95	1
5	Raina Sava Ramadhan	95	1
6	Riezka Nadya Putri	95	1
7	Aura Sabrina Putri	90	2
8	Tafakanara Alsajada D. A	90	2
9	Nuryan Husnul Khotimah	85	3

Jumlah kelompok atas (JA) = 9 orang

Kelompok Kelas Bawah (JB)

Diambil dari barisan siswa dengan skor terendah.

No	Nama	Skor	Soal Salah
1	Ahmad Faiz Raziq	80	4
2	Bagas Putra Kholik Pratama	80	4
3	M. Zaky Azwa Al Zevaro	80	4
4	Muhammad Akbar Mubarak	80	4

5	Sulthan Laksamana Agung	80	4
6	Rio Dwi Arviansyah	75	5
7	M. Datik Pratama D. P	70	6
8	Muhammad Fahri Alakhs	60	8
9	Muhammad Rakha Kiaky	60	8

Jumlah kelompok bawah (JB) = 9 orang

Tabel Hasil Analisis Daya Pembeda

No	BA	BB	JA	JB	PA	PB	DP (PA - PB)	Keterangan
1	9	4	9	9	1,00	0,44	0,56	Baik
2	9	5	9	9	1,00	0,56	0,44	Baik
3	9	3	9	9	1,00	0,33	0,67	Baik Sekali
4	8	4	9	9	0,89	0,44	0,45	Baik
5	7	6	9	9	0,78	0,67	0,11	Jelek
6	9	5	9	9	1,00	0,56	0,44	Baik
7	8	7	9	9	0,89	0,78	0,11	Jelek
8	9	4	9	9	1,00	0,44	0,56	Baik
9	8	5	9	9	0,89	0,56	0,33	Cukup
10	7	6	9	9	0,78	0,67	0,11	Jelek
11	6	5	9	9	0,67	0,56	0,11	Jelek
12	8	4	9	9	0,89	0,44	0,45	Baik
13	9	5	9	9	1,00	0,56	0,44	Baik
14	9	3	9	9	1,00	0,33	0,67	Baik Sekali
15	8	7	9	9	0,89	0,78	0,11	Jelek
16	9	5	9	9	1,00	0,56	0,44	Baik
17	8	4	9	9	0,89	0,44	0,45	Baik
18	9	4	9	9	1,00	0,44	0,56	Baik
19	7	6	9	9	0,78	0,67	0,11	Jelek
20	9	3	9	9	1,00	0,33	0,67	Baik Sekali

Kesimpulan:

Dapat disimpulkan bahwa instrumen tes ini memiliki kualitas daya pembeda yang bervariasi dengan mayoritas butir soal berada pada kategori Baik dan Baik Sekali. Dari total 20 butir soal, terdapat 3 soal (nomor 3, 14, dan 20) yang memiliki kualitas Baik Sekali dengan indeks sebesar 0,67, serta 10 butir soal berkategori Baik yang menunjukkan kemampuan efektif dalam membedakan antara kelompok siswa atas (JA) dan kelompok bawah (JB). Namun, masih ditemukan 6 butir soal berkategori Jelek (nomor 5, 7, 10, 11, 15, dan 19) serta 1 soal kategori Cukup yang memerlukan revisi atau penggantian agar instrumen evaluasi ini menjadi lebih akurat dalam mengukur kompetensi siswa secara keseluruhan.

B. Esai

No Soal	DP	Keterangan
1	0,48	Baik
2	0,45	Baik
3	0,53	Baik
4	0,42	Baik
5	0,36	Cukup

Kesimpulan:

Berdasarkan hasil analisis daya pembeda pada 5 butir soal uraian, diperoleh bahwa sebagian besar soal memiliki kemampuan yang baik dalam membedakan peserta didik berdasarkan tingkat penguasaan materi. Sebanyak 4 soal berada pada kategori baik dan 1 soal berada pada kategori cukup. Selain itu, tidak ditemukan soal yang berkategori jelek maupun negatif. Hasil tersebut menunjukkan bahwa instrumen uraian yang digunakan telah cukup efektif dalam mengidentifikasi perbedaan kemampuan antara peserta didik yang berprestasi tinggi dan peserta didik yang berprestasi rendah.

5. Distraktor (Pengecoh)

Sebuah pilihan jawaban dikatakan memiliki daya distraktor yang baik apabila:

- Dipilih minimal 5% peserta tes.
- Lebih banyak dipilih oleh kelompok bawah dibandingkan kelompok atas.
- Berfungsi sebagai pengecoh bagi peserta didik yang belum memahami materi.

Rumus:

$$D = \frac{A}{N} \times 100\%$$

Keterangan:

- D = Tingkat Distraktor (%)
- A = Jumlah siswa yang memilih opsi jawaban
- N = Jumlah seluruh siswa (18 siswa)

Tabel Interpretasi Distraktor

Besarnya D	Keterangan
$D \geq 5\%$	Diterima Baik
$5\% > D > 0$	Direvisi
$D = 0$	Ditolak/Buruk

Tabel Hasil Analisis Distraktor (Pengecoh)

No	A	B	C	D	E	Keterangan
1	0%	11%	89%	0%	0%	B diterima, C (Kunci) diterima, A ditolak, D ditolak, E ditolak karena tidak memenuhi syarat 5%
2	11%	0%	89%	0%	0%	A diterima, C (Kunci) diterima, B ditolak, D ditolak, E ditolak karena tidak memenuhi syarat 5%
3	0%	0%	100%	0%	0%	C (Kunci) diterima, A ditolak, B ditolak, D ditolak, E ditolak karena tidak memenuhi syarat 5%
4	6%	89%	6%	0%	0%	A diterima, B (Kunci) diterima, C diterima, D ditolak, E ditolak karena tidak memenuhi syarat 5%
5	0%	67%	22%	11%	0%	B (Kunci) diterima, C diterima, D diterima, A ditolak, E ditolak karena tidak memenuhi syarat 5%
6	0%	0%	94%	6%	0%	C (Kunci) diterima, D diterima, A ditolak, B ditolak, E ditolak karena tidak memenuhi syarat 5%
7	11%	0%	83%	6%	0%	A diterima, C (Kunci) diterima, D diterima, B ditolak, E ditolak karena tidak memenuhi syarat 5%
8	6%	6%	67%	22%	0%	A diterima, B diterima, C (Kunci) diterima, D diterima, E ditolak karena tidak memenuhi syarat 5%
9	11%	78%	6%	6%	0%	A diterima, B (Kunci) diterima, C diterima, D diterima, E ditolak karena tidak memenuhi syarat 5%
10	0%	11%	83%	6%	0%	B diterima, C (Kunci) diterima, D diterima, A ditolak, E ditolak karena tidak memenuhi syarat 5%
11	6%	6%	61%	28%	0%	A diterima, B diterima, C (Kunci) diterima, D diterima, E ditolak karena tidak memenuhi syarat 5%
12	11%	0%	83%	6%	0%	A diterima, C (Kunci) diterima, D diterima, B ditolak, E ditolak karena tidak memenuhi syarat 5%
13	0%	11%	89%	0%	0%	B diterima, C (Kunci) diterima, A ditolak, D ditolak, E ditolak karena tidak memenuhi syarat 5%
14	0%	6%	94%	0%	0%	B diterima, C (Kunci) diterima, A ditolak, D ditolak, E ditolak karena tidak memenuhi syarat 5%
15	11%	11%	78%	0%	0%	A diterima, B diterima, C (Kunci) diterima, D ditolak, E ditolak karena tidak memenuhi syarat 5%
16	0%	0%	89%	11%	0%	C (Kunci) diterima, D diterima, A ditolak, B ditolak, E ditolak karena tidak memenuhi syarat 5%
17	0%	0%	94%	6%	0%	C (Kunci) diterima, D diterima, A ditolak, B ditolak, E ditolak karena tidak memenuhi syarat 5%
18	6%	0%	89%	6%	0%	A diterima, C (Kunci) diterima, D diterima, B ditolak, E ditolak karena tidak memenuhi syarat 5%

19	11%	67%	17%	6%	0%	A diterima, B (Kunci) diterima, C diterima, D diterima, E ditolak karena tidak memenuhi syarat 5%
20	0%	6%	89%	6%	0%	B diterima, C (Kunci) diterima, D diterima, A ditolak, E ditolak karena tidak memenuhi syarat 5%

Kesimpulan:

Berdasarkan Tabel Hasil Analisis Distraktor (Pengecoh) pada 20 butir soal pilihan ganda, dapat disimpulkan bahwa fungsi pengecoh dalam instrumen tes ini berjalan cukup efektif namun memiliki variasi kualitas yang signifikan pada beberapa nomor. Sebagian besar soal memiliki opsi pengecoh yang Diterima karena berhasil menarik minat pilih dari minimal 5% peserta tes (setara dengan 1 orang siswa atau lebih), terutama pada soal nomor 4, 5, 8, 9, 11, dan 19 yang memiliki sebaran jawaban salah paling merata. Di sisi lain, pada soal-soal yang cenderung sangat mudah seperti nomor 3 dan 17, seluruh opsi pengecohnya Ditolak (0%) karena tidak ada satu pun siswa yang memilihnya. Secara keseluruhan, hasil ini menunjukkan bahwa sebagian besar opsi distraktor sudah mampu mengecoh siswa yang belum menguasai materi, meskipun butir soal dengan pengecoh yang ditolak tetap memerlukan revisi atau penggantian opsi jawaban agar kualitas instrumen tes di masa mendatang menjadi lebih seimbang dan objektif.

ANALISIS ASPEK KETERAMPILAN

Kategori	Jumlah	Persentase
Sangat Baik (A)	9	50,00%
Baik (B)	6	33,33%
Cukup (C)	3	16,67%
Perlu Bimbingan	0	00,00%
Total	18	100,00%

Keterangan:

Berdasarkan hasil penilaian aspek keterampilan terhadap 18 peserta didik, sebanyak 9 siswa (50,00%) memperoleh kategori Sangat Baik (A), 6 siswa (33,33%) memperoleh kategori Baik (B), dan 3 siswa (16,67%) memperoleh kategori Cukup (C). Dengan demikian, sebagian besar peserta didik telah menunjukkan keterampilan yang berada pada kategori Sangat Baik.

ANALISIS ASPEK SIKAP

Kategori	Jumlah	Persentase
Sangat Baik (A)	7	38,89%
Baik (B)	6	33,33%
Cukup (C)	5	27,78%
Perlu Bimbingan	0	00,00%
Total	18	100,00%

Keterangan:

Berdasarkan hasil penilaian aspek sikap terhadap 18 peserta didik, sebanyak 7 siswa (38,89%) memperoleh kategori Sangat Baik (A), 6 siswa (33,33%) memperoleh kategori Baik (B), dan 5 siswa (27,78%) memperoleh kategori Cukup (C). Hasil tersebut menunjukkan bahwa sebagian besar peserta didik telah menunjukkan sikap yang baik dalam proses pembelajaran, dengan dominasi pada kategori Sangat Baik dan Baik sebesar 72,22% dari seluruh peserta didik.