

## **ANALISIS KUALITAS TES DAN BUTIR SOAL**

Disusun Untuk Memenuhi Tugas Matakuliah Evaluasi Pembelajaran Ekonomi

### **Dosen Pengampu:**

Prof. Dr. Herpratiwi, M.Pd.

Dr. Pujiati, S.Pd., M. Pd.

Fanni Rahmawati, S.Pd., M.Pd.



Disusun Oleh:

Alzirah Sabrina

2413041049

**PROGRAM STUDI PENDIDIKAN EKONOMI  
JURUSAN PENDIDIKAN ILMU PENGETAHUAN SOSIAL  
FAKULTAS KEGURUAN DAN ILMU PENDIDIKAN UNIVERSITAS LAMPUNG  
2026**

## A. Pendahuluan

Evaluasi pembelajaran merupakan bagian penting dalam proses pendidikan karena digunakan untuk mengetahui sejauh mana peserta didik memahami materi dan mencapai kompetensi yang telah ditetapkan setelah mengikuti kegiatan pembelajaran. Dalam pelaksanaannya, instrumen evaluasi yang digunakan tidak hanya sekadar dibuat, tetapi juga perlu dianalisis kualitasnya agar hasil yang diperoleh benar-benar dapat dipercaya dan dijadikan dasar dalam pengambilan keputusan pembelajaran. Analisis kualitas tes dan butir soal dilakukan untuk mengetahui apakah soal yang digunakan sudah mampu mengukur kemampuan peserta didik secara tepat, konsisten, objektif, dan sesuai dengan tujuan pembelajaran.

Selain itu, melalui analisis ini dapat diketahui tingkat kesukaran soal, kemampuan soal dalam membedakan peserta didik berkemampuan tinggi dan rendah, serta efektivitas pilihan jawaban pengecoh pada soal pilihan ganda. Pada penelitian ini, instrumen evaluasi yang digunakan terdiri atas 20 soal pilihan ganda dan 5 soal esai dengan jumlah responden sebanyak 35 peserta didik kelas XI IPA A.3 MAN 1 Metro pada materi Pendapatan Nasional mata pelajaran Ekonomi. Adapun aspek yang dianalisis meliputi validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas distraktor. Dengan dilakukannya analisis kualitas tes dan butir soal, diharapkan instrumen evaluasi yang digunakan dapat mengukur kemampuan peserta didik secara lebih akurat sehingga hasil evaluasi dapat dimanfaatkan sebagai bahan refleksi dan perbaikan dalam proses pembelajaran.

## B. Uji Validitas

Validitas merupakan tingkat ketepatan suatu instrumen dalam mengukur kemampuan peserta didik sesuai indikator pembelajaran. Uji validitas dilakukan menggunakan aplikasi SPSS dengan teknik *Product Moment Pearson* berdasarkan skor jawaban peserta didik. Soal dinyatakan valid apabila  $r$  hitung  $>$   $r$  tabel. Hasil uji validitas soal pilihan ganda disajikan pada tabel berikut.

1. Pilihan Ganda			
No.	r Hitung	r Tabel	Keterangan
1.	- (tidak dapat dihitung)	0,334	Tidak dapat diuji
2.	0,679	0,334	Valid
3.	0,616	0,334	Valid
4.	0,501	0,334	Valid
5.	0,522	0,334	Valid
6.	0,743	0,334	Valid
7.	0,502	0,334	Valid
8.	0,893	0,334	Valid
9.	0,893	0,334	Valid
10.	0,823	0,334	Valid

11.	0,893	0,334	Valid
12.	0,734	0,334	Valid
13.	0,743	0,334	Valid
14.	0,596	0,334	Valid
15.	0,893	0,334	Valid
16.	0,781	0,334	Valid
17.	0,502	0,334	Valid
18.	0,893	0,334	Valid
19.	0,893	0,334	Valid
20.	0,743	0,334	Valid

#### Interpretasi Uji Validitas Soal Pilihan Ganda

Berdasarkan hasil uji validitas menggunakan SPSS, sebanyak 19 soal pilihan ganda dinyatakan valid karena nilai  $r$  hitung lebih besar dari  $r$  tabel (0,334), sedangkan 1 soal tidak dapat diuji karena seluruh responden memberikan jawaban yang sama sehingga korelasi tidak dapat dihitung. Hal ini menunjukkan bahwa soal telah sesuai dengan indikator pembelajaran dan layak digunakan sebagai instrumen penilaian.

2. Esai			
No.	$r$ Hitung	$r$ Tabel	Keterangan
1.	0,937	0,334	Valid
2.	0,887	0,334	Valid
3.	0,896	0,334	Valid
4.	0,965	0,334	Valid
5.	0,939	0,334	Valid

#### Interpretasi Uji Validitas Soal Esai

Berdasarkan hasil uji validitas menggunakan SPSS, seluruh 5 soal esai dinyatakan valid karena nilai  $r$  hitung lebih besar dari  $r$  tabel. Hal ini menunjukkan bahwa soal esai telah sesuai dengan indikator pembelajaran.

3. Keterampilan			
No.	$r$ Hitung	$r$ Tabel	Keterangan
1.	0,854	0,334	Valid
2.	0,690	0,334	Valid
3.	0,701	0,334	Valid
4.	0,601	0,334	Valid
5.	0,789	0,334	Valid

4. Sikap			
No.	$r$ Hitung	$r$ Tabel	Keterangan
1.	0,750	0,334	Valid

2.	0,621	0,334	Valid
3.	0,738	0,334	Valid
4.	0,493	0,334	Valid
5.	0,828	0,334	Valid

#### Interpretasi Uji Validitas Aspek Keterampilan dan Sikap

Berdasarkan hasil uji validitas menggunakan bantuan IBM SPSS Statistics, seluruh 5 butir instrumen aspek sikap dan 5 butir instrumen aspek keterampilan dinyatakan valid karena memiliki nilai  $r$  hitung lebih besar daripada  $r$  tabel (0,334). Hal ini menunjukkan bahwa setiap butir instrumen telah sesuai dengan indikator pembelajaran, mampu mengukur aspek yang hendak dinilai, serta layak digunakan sebagai instrumen penilaian dalam proses evaluasi pembelajaran.

### C. Uji Reliabilitas

Reliabilitas digunakan untuk mengetahui tingkat konsistensi suatu instrumen evaluasi. Instrumen yang reliabel akan memberikan hasil yang relatif tetap apabila digunakan dalam kondisi yang sama. Uji reliabilitas dilakukan menggunakan metode Cronbach Alpha dengan bantuan aplikasi SPSS. Kriteria reliabilitas yaitu  $\alpha > 0,90$  = sangat tinggi,  $\alpha > 0,70$  = tinggi/reliabel,  $\alpha > 0,50$  = cukup, dan  $\alpha < 0,50$  = rendah.

Aspek	Cronbach's Alpha	Keterangan
Pengetahuan (Pilihan Ganda)	0,939	Sangat Reliabel
Pengetahuan (Esai)	0,953	Sangat Reliabel
Keterampilan	0,769	Reliabel
Sikap	0,720	Reliabel

Berdasarkan hasil uji reliabilitas menggunakan bantuan IBM SPSS Statistics, instrumen aspek pengetahuan berupa soal pilihan ganda memperoleh nilai Cronbach's Alpha sebesar 0,939, sedangkan instrumen soal esai memperoleh nilai Cronbach's Alpha sebesar 0,953. Kedua nilai tersebut menunjukkan bahwa instrumen evaluasi memiliki tingkat reliabilitas sangat tinggi, sehingga layak digunakan sebagai alat evaluasi pembelajaran.

### D. Analisis Tingkat Kesukaran Soal

Tingkat kesukaran soal digunakan untuk mengetahui apakah suatu soal tergolong mudah, sedang, atau sulit bagi peserta didik. Analisis ini dilakukan dengan membandingkan jumlah siswa yang menjawab benar dengan jumlah seluruh siswa menggunakan rumus  $P = B/N$ , dimana  $P$  adalah indeks kesukaran,  $B$  jumlah siswa menjawab benar, dan  $N$  jumlah seluruh siswa. Kriteria tingkat kesukaran yaitu 0,00–0,30 = sukar, 0,31–0,70 = sedang, dan 0,71–1,00 = mudah. Berdasarkan hasil pengerjaan siswa, sebagian besar soal berada pada kategori mudah hingga sedang. Soal LOTS cenderung lebih mudah, sedangkan soal MOTS dan HOTS lebih sulit karena membutuhkan kemampuan analisis dan penerapan konsep ekonomi dalam kehidupan sehari-hari.

1. Tabel Analisis Tingkat Kesukaran Soal Pilihan Ganda

No.	Benar	Salah	TK	Kategori
1	35	0	1,00	Mudah
2	28	7	0,80	Mudah
3	33	2	0,94	Mudah
4	27	8	0,77	Mudah
5	32	3	0,91	Mudah
6	34	1	0,97	Mudah
7	34	1	0,97	Mudah
8	33	2	0,94	Mudah
9	33	2	0,94	Mudah
10	31	4	0,89	Mudah
11	33	2	0,94	Mudah
12	29	6	0,83	Mudah
13	34	1	0,97	Mudah
14	31	4	0,89	Mudah
15	33	2	0,94	Mudah
16	30	5	0,86	Mudah
17	34	1	0,97	Mudah
18	33	2	0,94	Mudah
19	33	2	0,94	Mudah
20	34	1	0,97	Mudah

Berdasarkan hasil analisis tingkat kesukaran, seluruh soal berada pada kategori mudah sehingga sesuai dengan kemampuan peserta didik.

2. Tabel Analisis Tingkat Kesukaran Soal Esai

Soal	Jumlah Skor	Rata-rata	TK	Kategori
S1	580	16,57	0,83	Mudah
S2	550	15,71	0,79	Mudah
S3	555	15,86	0,79	Mudah
S4	535	15,29	0,76	Mudah
S5	600	17,14	0,86	Mudah

Berdasarkan hasil analisis tingkat kesukaran, seluruh soal berada pada kategori mudah sehingga sesuai dengan kemampuan peserta didik.

3. Tabel Analisis Tingkat Kesukaran Aspek Keterampilan

Item	Jumlah Skor	Rata-rata	SMI	TK	Kategori
------	-------------	-----------	-----	----	----------

K1	102	2,91	4	0,73	Mudah
K2	96	2,74	4	0,69	Sedang
K3	100	2,86	4	0,72	Mudah
K4	102	2,91	4	0,73	Mudah
K5	105	3,00	4	0,75	Mudah

Berdasarkan hasil uji tingkat kesukaran instrumen aspek keterampilan, diperoleh bahwa sebagian besar item berada pada kategori mudah, sedangkan satu item berada pada kategori sedang. Hasil ini menunjukkan bahwa indikator pada instrumen keterampilan relatif mudah dipahami dan dipenuhi oleh peserta didik, sehingga instrumen layak digunakan dalam proses penilaian pembelajaran.

#### 4. Tabel Analisis Tingkat Kesukaran Aspek Keterampilan

Item	Jumlah Skor	Rata-rata	SMI	TK	Kategori
S1	109	3,11	4	0,78	Mudah
S2	125	3,57	4	0,89	Mudah
S3	118	3,37	4	0,84	Mudah
S4	114	3,26	4	0,81	Mudah
S5	116	3,31	4	0,83	Mudah

Berdasarkan hasil uji tingkat kesukaran, seluruh item pada instrumen aspek sikap termasuk dalam kategori mudah. Hal ini menunjukkan bahwa indikator yang digunakan dapat dicapai dengan baik oleh peserta didik dan instrumen layak digunakan sebagai alat penilaian aspek sikap.

### E. Analisis Daya Pembeda

Daya pembeda merupakan kemampuan soal dalam membedakan peserta didik berkemampuan tinggi dan rendah. Analisis dilakukan menggunakan rumus  $D = (BA - BB)/N$ , dimana BA adalah jumlah jawaban benar kelompok atas, BB jumlah jawaban benar kelompok bawah, dan N jumlah seluruh siswa. Kriteria daya pembeda yaitu  $D < 0,20$  = jelek,  $0,20-0,40$  = cukup,  $0,41-0,70$  = baik, dan  $>0,70$  = sangat baik. Berdasarkan hasil analisis, sebagian besar soal memiliki daya pembeda cukup hingga baik sehingga mampu membedakan kemampuan peserta didik. Jumlah siswa sebanyak 35 orang dibagi menjadi kelompok atas 17 siswa dan kelompok bawah 17 siswa menggunakan metode 50%.

#### 1. Tabel Analisis Daya Pembeda Soal Pilihan Ganda

No	BA	BB	DP	Kategori
1	17	17	0,00	Jelek
2	17	11	0,35	Cukup
3	17	15	0,12	Jelek
4	17	10	0,41	Baik
5	17	14	0,18	Jelek

6	17	16	0,06	Jelek
7	17	16	0,06	Jelek
8	17	15	0,12	Jelek
9	17	15	0,12	Jelek
10	17	13	0,24	Cukup
11	17	15	0,12	Jelek
12	17	11	0,35	Cukup
13	17	16	0,06	Jelek
14	17	13	0,24	Cukup
15	17	15	0,12	Jelek
16	17	12	0,29	Cukup
17	17	16	0,06	Jelek
18	17	15	0,12	Jelek
19	17	15	0,12	Jelek
20	17	16	0,06	Jelek

Berdasarkan hasil analisis daya pembeda menggunakan metode 50%, diperoleh 1 soal kategori baik, 5 soal kategori cukup, dan 14 soal kategori jelek. Hasil tersebut menunjukkan bahwa sebagian besar butir soal masih perlu diperbaiki agar lebih mampu membedakan peserta didik berkemampuan tinggi dan rendah.

## 2. Tabel Analisis Daya Pembeda Soal Esai

Soal	Mean Kelompok Atas	Mean Kelompok Bawah	DP	Kategori
E1	19,12	12,35	0,34	Cukup
E2	18,82	11,76	0,35	Cukup
E3	18,53	12,06	0,32	Cukup
E4	18,24	10,88	0,37	Cukup
E5	19,41	13,24	0,31	Cukup

Berdasarkan hasil analisis daya pembeda soal esai, seluruh 5 butir soal berada pada kategori cukup. Hal ini menunjukkan bahwa soal esai sudah cukup mampu membedakan peserta didik yang memiliki kemampuan tinggi dan rendah.

## 3. Tabel Analisis Daya Pembeda Aspek Keterampilan dan Sikap

### a. Aspek Keterampilan

Item	Mean Kelompok Atas	Mean Kelompok Bawah	DP	Kategori
K1	3,41	2,35	0,27	Cukup
K2	3,29	2,06	0,31	Cukup
K3	3,47	2,18	0,32	Cukup
K4	3,53	2,12	0,35	Cukup
K5	3,59	2,24	0,34	Cukup

### b. Aspek Sikap

Item	Mean Atas	Mean Bawah	DP	Kategori
S1	3,41	2,24	0,29	Cukup
S2	3,53	2,18	0,34	Cukup
S3	3,47	2,12	0,34	Cukup
S4	3,35	2,06	0,32	Cukup
S5	3,59	2,29	0,33	Cukup

Berdasarkan hasil uji daya pembeda, seluruh item pada instrumen aspek sikap dan aspek keterampilan berada pada kategori cukup. Hal ini menunjukkan bahwa setiap butir instrumen telah memiliki kemampuan yang cukup dalam membedakan peserta didik dengan tingkat sikap dan keterampilan tinggi maupun rendah, sehingga instrumen layak digunakan dalam proses penilaian pembelajaran.

#### F. Analisis Distraktor

Distraktor atau pengecoh merupakan pilihan jawaban salah pada soal pilihan ganda yang digunakan untuk mengetahui apakah peserta didik benar-benar memahami materi. Analisis distraktor dilakukan untuk melihat efektivitas pilihan jawaban yang disusun. Berdasarkan hasil analisis, sebagian besar distraktor sudah berfungsi dengan baik karena dipilih oleh beberapa peserta didik sehingga pilihan jawaban dinilai cukup logis dan tidak mudah ditebak.

Tabel Analisis Distraktor Soal Pilihan Ganda

No	Kunci	A	B	C	D	E	Analisis Distraktor
1	B	0	35	0	0	0	Distraktor A,C,D,E tidak berfungsi
2	C	7	1	26	0	1	Distraktor A berfungsi, B,D,E tidak berfungsi
3	B	1	34	0	0	0	Distraktor A,C,D,E tidak berfungsi
4	D	4	2	2	27	0	Distraktor A,B,C berfungsi, E tidak berfungsi
5	A	34	1	0	0	0	Distraktor B,C,D,E tidak berfungsi
6	B	1	34	0	0	0	Distraktor A,C,D,E tidak berfungsi
7	B	0	34	1	0	0	Distraktor A,C,D,E tidak berfungsi
8	A	33	2	0	0	0	Distraktor B berfungsi, C,D,E tidak berfungsi
9	D	0	2	0	33	0	Distraktor B berfungsi, A,C,E tidak berfungsi
10	C	3	1	31	0	0	Distraktor A berfungsi, B,D,E tidak berfungsi
11	B	1	33	0	1	0	Distraktor A,C,D,E tidak berfungsi
12	A	29	1	1	4	0	Distraktor D berfungsi, B,C,E tidak berfungsi
13	B	0	35	0	0	0	Distraktor A,C,D,E tidak berfungsi
14	C	2	1	31	1	0	Distraktor A berfungsi, B,D,E tidak berfungsi
15	B	0	33	2	0	0	Distraktor C berfungsi, A,D,E tidak berfungsi
16	B	3	30	2	0	0	Distraktor A,C berfungsi, D,E tidak berfungsi

17	A	34	1	0	0	0	Distraktor B,C,D,E tidak berfungsi
18	B	0	33	2	0	0	Distraktor C berfungsi, A,D,E tidak berfungsi
19	B	2	33	0	0	0	Distraktor A berfungsi, C,D,E tidak berfungsi
20	A	34	1	0	0	0	Distraktor B,C,D,E tidak berfungsi

Berdasarkan hasil analisis distraktor, hanya beberapa distraktor yang berfungsi dengan baik karena dipilih oleh minimal 5% peserta didik, sedangkan sebagian besar distraktor masih belum efektif karena jarang dipilih atau tidak dipilih sama sekali.

### G. Kesimpulan

Berdasarkan hasil evaluasi terhadap kualitas tes dan soal mengenai Pendapatan Nasional, secara umum, alat ukur yang digunakan telah memenuhi standar instrumen yang berkualitas. Uji validitas menunjukkan bahwa sebanyak 19 soal pilihan ganda dinyatakan valid, sedangkan 1 soal tidak dapat diuji karena seluruh responden memberikan jawaban yang sama. Seluruh soal esai, instrumen keterampilan, dan instrumen sikap dinyatakan valid. Selain itu, uji reliabilitas juga menunjukkan nilai Cronbach Alpha yang sangat tinggi, yang berarti instrumen tersebut dinilai stabil dan cocok untuk dipakai dalam evaluasi pembelajaran.

Dalam analisis tingkat kesulitan, sebagian besar soal tergolong mudah sehingga sesuai dengan level kemampuan siswa. Soal yang disusun juga mencakup berbagai tingkat kognitif LOTS, MOTS, dan HOTS, sehingga tidak hanya mengevaluasi kemampuan mengingat, tetapi juga kemampuan memahami, menghitung, dan menganalisis masalah ekonomi dalam kehidupan sehari-hari.

Analisis daya pembeda menunjukkan bahwa ada beberapa soal yang termasuk dalam kategori buruk sehingga perlu perbaikan agar soal-soal tersebut dapat lebih efektif dalam membedakan siswa dengan kemampuan tinggi dan rendah. Di sisi lain, analisis distraktor menunjukkan bahwa hanya sebagian kecil distraktor berfungsi efektif, sementara sebagian besar opsi pengecoh masih perlu diperbaiki agar lebih mampu menarik peserta didik yang belum memahami materi.

Secara keseluruhan, alat ukur evaluasi yang digunakan sudah tergolong baik dan dapat diterapkan dalam proses pembelajaran ekonomi. Namun, beberapa soal masih memerlukan revisi, terutama terkait dengan daya pembeda dan distraktor agar kualitas instrumen dapat ditingkatkan untuk penggunaan di masa mendatang. Dengan adanya analisis kualitas tes dan soal, guru dapat memahami kelebihan dan kekurangan dari instrumen evaluasi, sehingga hasil penilaian dapat dijadikan acuan untuk memperbaiki proses pembelajaran dengan lebih efektif dan profesional.