

ANALISIS BUTIR SOAL

Aspek Pengetahuan (Kognitif) Pilihan Ganda

Nama	Ratih Apriyani
NPM	2413031073
Kelas	24 C
Mata Kuliah	Evaluasi Pembelajaran Ekonomi
Sekolah	SMA Negeri 6 Metro
Kelas Observasi	X.6
Materi	Perbankan
Jumlah Soal PG	20 butir (opsi A–E)
Jumlah Responden	25 siswa/i

A. Pendahuluan

Dokumen ini menyajikan analisis butir soal pilihan ganda dari hasil observasi evaluasi pembelajaran Ekonomi kelas X.6 SMA Negeri 6 Metro dengan materi Perbankan, sebagai kelengkapan dari hasil observasi yang telah disusun oleh Ratih Apriyani. Data jawaban diperoleh dari rekapitulasi Google Form (file respons mentah) yang diisi oleh 25 peserta didik. Tes terdiri dari 20 soal pilihan ganda dengan opsi A–E.

Karena dokumen hasil observasi awal hanya memuat skor total per siswa (bukan jawaban per item), kunci jawaban setiap soal direkonstruksi dari pola jawaban siswa yang memperoleh skor sempurna (20/20), kemudian divalidasi ulang terhadap skor seluruh 25 siswa. Hasil validasi menunjukkan kecocokan sempurna (0 dari 25 siswa mengalami selisih skor), sehingga kunci jawaban yang digunakan dalam analisis ini dapat dipertanggungjawabkan secara matematis.

Aspek yang dianalisis meliputi: Validitas (korelasi point-biserial), Reliabilitas (KR-20), Tingkat Kesukaran (P), Daya Pembeda (D), dan Efektivitas Distraktor, mengikuti kerangka analisis yang sama dengan dokumen acuan “Analisis Butir Soal” milik Silviana Febriani.

B. Data Peserta dan Hasil Tes

Jumlah peserta tes: 25 siswa/i kelas X.6 SMA Negeri 6 Metro. Skor pilihan ganda berikut diambil dari kolom "Score" pada rekap Google Form (skala 0–100, dengan 5 poin per soal benar).

No	Nama Peserta Didik	Skor PG	Jml Benar	% Benar
1	chika almira	100	20	100%
2	syifa mutia rahma	100	20	100%
3	Aira syafa abeluna	100	20	100%
4	dinda khania	100	20	100%
5	SINTA KARUNIA MEGA	100	20	100%
6	SHIRTON FATAH ATHOILLAH	75	15	75%
7	VIOLA AGUSTIN	100	20	100%
8	Ainur Maulana Bayu Atmaja	100	20	100%
9	Lussal bara elcco	50	10	50%
10	Azka princestia	55	11	55%
11	RAYENDRA FEBRIANSYAH	100	20	100%
12	Sri alysa handayani	85	17	85%
13	ASHIF AZZAM SADIKIN	85	17	85%
14	BINTANG CITRA DEWI	80	16	80%
15	Putri Adonia Yessica Andriyana	75	15	75%
16	Jovanka Amelia Putri	60	12	60%
17	rahma mubaroka	90	18	90%
18	ghista dina maula	90	18	90%

No	Nama Peserta Didik	Skor PG	Jml Benar	% Benar
19	Tiara Alisa Agustina	80	16	80%
20	latifa hummayroh	95	19	95%
21	AHMAD HANIF	100	20	100%
22	alchanu exfal malatta	85	17	85%
23	IMAM NABIL AL FAJRI	100	20	100%
24	Ras'sya aditiya pratama	100	20	100%
25	Ericka Mega Aulia	85	17	85%

C. Analisis Butir Soal Pilihan Ganda

1. Validitas

Validitas diukur menggunakan korelasi point-biserial (r_{pb}). Soal dinyatakan valid apabila $r_{pb} \geq r_{tabel}$. Dengan $N = 25$ dan taraf signifikansi $\alpha = 0,05$ (two-tailed, $df = N-2 = 23$), nilai $r_{tabel} = 0.396$ [dihitung langsung dari distribusi t, bukan dikutip dari tabel cetak].

No. Soal	r_{pb}	r_{tabel}	Keterangan
1	0.351	0.396	Tidak Valid
2	0.152	0.396	Tidak Valid
3	0.796	0.396	Valid
4	0.701	0.396	Valid
5	0.59	0.396	Valid
6	0.52	0.396	Valid
7	0.63	0.396	Valid

No. Soal	r_pb	r_tabel	Keterangan
8	0.381	0.396	Tidak Valid
9	0.765	0.396	Valid
10	0.401	0.396	Valid
11	0.273	0.396	Tidak Valid
12	0.174	0.396	Tidak Valid
13	0.701	0.396	Valid
14	NaN	0.396	Tidak Valid
15	0.651	0.396	Valid
16	0.765	0.396	Valid
17	NaN	0.396	Tidak Valid
18	0.557	0.396	Valid
19	0.105	0.396	Tidak Valid
20	NaN	0.396	Tidak Valid

Keterangan: Soal valid = 11 butir, Soal tidak valid = 9 butir dari total 20 soal.

2. Reliabilitas

Reliabilitas dihitung menggunakan rumus Kuder-Richardson 20 (KR-20). Tes dinyatakan reliabel jika $KR-20 \geq 0,70$ (kategori tinggi).

Komponen	Nilai
Jumlah Soal (k)	20
Jumlah Peserta Didik (N)	25
Rata-rata Skor (M)	17.52
Varians Skor (S^2)	8.73

Komponen	Nilai
ΣpQ	1.878
Koefisien KR-20	0.8261
Kategori	Tinggi

Kesimpulan: Koefisien reliabilitas KR-20 = 0.8261, artinya tes ini memiliki konsistensi internal yang tinggi dan layak digunakan sebagai instrumen penilaian dari sisi konsistensi. Namun, kategori reliabilitas tinggi ini perlu dibaca berdampingan dengan hasil validitas dan daya pembeda di bawah reliabilitas tinggi tidak menjamin setiap butir soal valid atau mampu membedakan kemampuan siswa.

3. Tingkat Kesukaran

Tingkat kesukaran (P) dihitung dengan rumus $P = \Sigma B / N$. Kategori: $P < 0,30$ = Sukar; $0,31-0,70$ = Sedang; $P > 0,70$ = Mudah.

No. Soal	ΣB	N	P	Kategori
1	23	25	0.92	Mudah
2	23	25	0.92	Mudah
3	19	25	0.76	Mudah
4	19	25	0.76	Mudah
5	12	25	0.48	Sedang
6	24	25	0.96	Mudah
7	20	25	0.8	Mudah
8	24	25	0.96	Mudah
9	19	25	0.76	Mudah
10	23	25	0.92	Mudah

No. Soal	ΣB	N	P	Kategori
11	22	25	0.88	Mudah
12	24	25	0.96	Mudah
13	23	25	0.92	Mudah
14	25	25	1	Mudah
15	23	25	0.92	Mudah
16	20	25	0.8	Mudah
17	25	25	1	Mudah
18	21	25	0.84	Mudah
19	24	25	0.96	Mudah
20	25	25	1	Mudah

Keterangan: Mudah = 19 soal (95%), Sedang = 1 soal (5%), Sukar = 0 soal (0%).

Distribusi tingkat kesukaran sangat tidak seimbang: hampir seluruh soal (19 dari 20) tergolong mudah. Soal dengan P mendekati atau sama dengan 1,00 perlu ditinjau ulang karena tidak mampu membedakan siswa berkemampuan tinggi dan rendah hal ini konsisten dengan banyaknya butir yang “Ditolak/Buang” pada hasil daya pembeda di bagian berikutnya.

4. Daya Pembeda

Daya pembeda (D) dihitung dengan rumus $D = P_a - P_b$. Kelompok atas dan bawah dibagi berdasarkan skor total, mengikuti proporsi yang sebanding dengan dokumen acuan: kelompok atas $n = 13$ siswa, kelompok bawah $n = 12$ siswa.

(Kelompok Atas)

No	Nama Siswa	Skor
1	chika almira	20

No	Nama Siswa	Skor
2	syifa mutia rahma	20
3	Aira syafa abeluna	20
4	dinda khania	20
5	SINTA KARUNIA MEGA	20
6	VIOLA AGUSTIN	20
7	Ainur Maulana Bayu Atmaja	20
8	Ras'sya aditiya pratama	20
9	RAYENDRA FEBRIANSYAH	20
10	IMAM NABIL AL FAJRI	20
11	AHMAD HANIF	20
12	latifa hummayroh	19
13	rahma mubaroka	18

(Kelompok Bawah)

No	Nama Siswa	Skor
1	ghista dina maula	18
2	alchanu exfal malatta	17
3	Sri alysa handayani	17
4	ASHIF AZZAM SADIKIN	17
5	Ericka Mega Aulia	17
6	BINTANG CITRA DEWI	16
7	Tiara Alisa Agustina	16

No	Nama Siswa	Skor
8	SHIRTON FATAH ATHOILLAH	15
9	Putri Adonia Yessica Andriyana	15
10	Jovanka Amelia Putri	12
11	Azka princestia	11
12	Lussal bara elcco	10

Rumus Daya Pembeda: $D = Pa - Pb$, dengan $Pa = Ba/Ja$ dan $Pb = Bb/Jb$.

Kategori: $D \geq 0,40$ = Sangat Bagus; $0,30-0,39$ = Cukup Baik; $0,20-0,29$ = Perlu Perbaikan; $D < 0,20$ = Ditolak/Buang.

No	Ba	Bb	Pa	Pb	D	Kategori
1	13	10	1	0.83	0.17	Ditolak/Buang
2	13	10	1	0.83	0.17	Ditolak/Buang
3	13	6	1	0.5	0.5	Sangat Bagus
4	13	6	1	0.5	0.5	Sangat Bagus
5	11	1	0.85	0.08	0.76	Sangat Bagus
6	13	11	1	0.92	0.08	Ditolak/Buang
7	13	7	1	0.58	0.42	Sangat Bagus
8	13	11	1	0.92	0.08	Ditolak/Buang
9	13	6	1	0.5	0.5	Sangat Bagus
10	13	10	1	0.83	0.17	Ditolak/Buang
11	12	10	0.92	0.83	0.09	Ditolak/Buang
12	13	11	1	0.92	0.08	Ditolak/Buang

No	Ba	Bb	Pa	Pb	D	Kategori
13	13	10	1	0.83	0.17	Ditolak/Buang
14	13	12	1	1	0	Ditolak/Buang
15	13	10	1	0.83	0.17	Ditolak/Buang
16	13	7	1	0.58	0.42	Sangat Bagus
17	13	12	1	1	0	Ditolak/Buang
18	13	8	1	0.67	0.33	Cukup Baik
19	13	11	1	0.92	0.08	Ditolak/Buang
20	13	12	1	1	0	Ditolak/Buang

Ringkasan Hasil Daya Pembeda

Kategori	Jumlah Soal
Sangat Bagus	6
Cukup Baik	1
Perlu Perbaikan	0
Ditolak/Buang	13

Analisis: Berdasarkan hasil analisis daya pembeda, terdapat 6 butir soal dengan daya pembeda sangat bagus, yaitu nomor 3, 4, 5, 7, 9, 16. Terdapat 1 butir soal kategori cukup baik, yaitu nomor 18. Tidak ada soal pada kategori perlu perbaikan. Sebanyak 13 butir soal (nomor 1, 2, 6, 8, 10, 11, 12, 13, 14, 15, 17, 19, 20) termasuk kategori ditolak/dibuang karena daya pembeda rendah, sebagian besar disebabkan oleh tingkat kesukaran yang terlalu mudah (P mendekati 1,00) sehingga hampir seluruh siswa baik kelompok atas maupun bawah menjawab benar. Secara proporsi, butir soal yang perlu direvisi (13 dari 20, atau 65%) lebih banyak dibanding butir yang layak dipertahankan.

5. Distraktor

Distraktor dinyatakan berfungsi jika dipilih oleh minimal 1 siswa dari N=25. Kategori berdasarkan jumlah pengecoh yang berfungsi: 4 = Sangat Baik; 3 = Baik; 2 = Cukup Baik; 1 = Kurang Baik; 0 = Tidak Baik.

No.	A	B	C	D	E	Kunci	Efektivitas
1	0	23	0	2	0	B	Kurang Baik
2	1	1	23	0	0	C	Cukup Baik
3	19	1	3	0	2	A	Baik
4	19	3	0	2	1	A	Baik
5	12	2	2	0	9	A	Baik
6	24	0	0	0	1	A	Kurang Baik
7	2	2	0	20	1	D	Baik
8	0	24	1	0	0	B	Kurang Baik
9	4	0	19	2	0	C	Cukup Baik
10	2	0	23	0	0	C	Kurang Baik
11	22	1	1	1	0	A	Baik
12	0	1	0	24	0	D	Kurang Baik
13	0	23	2	0	0	B	Kurang Baik
14	0	25	0	0	0	B	Tidak Baik
15	0	0	23	1	1	C	Cukup Baik
16	5	20	0	0	0	B	Kurang Baik
17	0	25	0	0	0	B	Tidak Baik
18	0	3	21	0	1	C	Cukup Baik
19	0	1	24	0	0	C	Kurang Baik
20	0	25	0	0	0	B	Tidak Baik

Ringkasan Efektivitas Distraktor

Kategori	Jumlah Soal
Sangat Baik	0
Baik	5
Cukup Baik	4
Kurang Baik	8
Tidak Baik	3

Keterangan: Sangat Baik = 0 soal, Baik = 5 soal, Cukup Baik = 4 soal, Kurang Baik = 8 soal, Tidak Baik = 3 soal. Tidak ada satu pun soal yang seluruh distraktornya berfungsi optimal (kategori Sangat Baik). Soal dengan efektivitas distraktor rendah (Kurang Baik/Tidak Baik) perlu direvisi dengan mengganti pengecoh agar lebih menarik dan logis bagi siswa, terutama pada soal-soal yang opsi pengecohnya sama sekali tidak dipilih (0 siswa).

D. Kesimpulan dan Rekomendasi

- Validitas: dari 20 soal, hanya 11 butir (55%) dinyatakan valid ($r_{pb} \geq 0.396$); 9 butir (45%) tidak valid dan memerlukan revisi atau penggantian.
- Reliabilitas: koefisien KR-20 = 0.8261 (kategori Tinggi), menunjukkan konsistensi internal instrumen secara keseluruhan baik, meskipun banyak butir individual bermasalah.
- Tingkat kesukaran: distribusi sangat tidak ideal 19 dari 20 soal (95%) tergolong mudah, hanya 1 soal sedang, dan tidak ada soal sukar. Instrumen ini tidak cukup menantang untuk membedakan level pemahaman siswa.
- Daya pembeda: 13 dari 20 soal (65%) masuk kategori ditolak/buang karena tingkat kesukaran yang terlalu mudah menghapus variasi antara kelompok atas dan bawah.

- Efektivitas distraktor: sebagian besar soal (11 dari 20) memiliki pengecoh yang tidak berfungsi optimal, karena opsi jawaban yang terlalu mudah ditebak.

Secara umum, instrumen soal pilihan ganda ini memiliki reliabilitas yang baik dari sisi konsistensi skor, namun kualitas butir soal secara individual masih rendah: mayoritas soal terlalu mudah, kurang valid, dan kurang mampu membedakan kemampuan siswa. Disarankan untuk merevisi pengecoh pada soal-soal dengan distraktor tidak berfungsi, serta menambah variasi tingkat kesukaran (sedang dan sukar) pada revisi instrumen berikutnya agar instrumen lebih representatif dalam mengukur capaian belajar siswa.